Does promoting one behavior distract from others? Evidence from a field experiment *

Hannah Trachtman †

January 24, 2022

Abstract

Impact evaluations of behavioral interventions typically focus on target outcomes. However, if effort and attention are costly, generic interventions might induce negative spillovers on other behaviors. I run a large field experiment in which individuals receive combinations of messages and incentives promoting two behaviors, meditation and meal logging. I find that the interventions reduce completion rates of the opposite behavior by 19-29%. Decomposing spillovers using a new taxonomy, I find that spillovers act in part as a fixed cost: interventions with larger target effects will not necessarily generate larger negative spillovers. I discuss implications for cost-effectiveness analysis.

JEL classification: D62, D91, I12

*I am so grateful to my dissertation committee for their help at every stage of this project: Dean Karlan, Chris Udry, Costas Meghir, and Jason Abaluck. For helpful discussions and comments, I thank Joseph Altonji, Gaurav Chiplunkar, Taha Choukhmane, Marvin Chun, Ori Heffetz, John Eric Humphries, Ro'ee Levy, Cormac McCarthy, Mushfiq Mobarak, Mark Rosenzweig, and Jaya Wen. For excellent research assistance I thank Osman Tuunteeyah. I would like to thank the Yale Economics Department and Economic Growth Center, the Sylff Foundation, the National Science Foundation, the Russell Sage Foundation, and the Institution for Social and Policy Studies for their generous support. This research was conducted under full IRB approval. All errors are my own.

†Department of Economics, The Hebrew University of Jerusalem, Mt. Scopus, Jerusalem, 9190501, Israel. hannah.trachtman@mail.huji.ac.il

1 Introduction

Seemingly small changes in behaviors like routine childhood immunization, healthy eating, and the use of clean cookstoves—among others—can lead to big economic benefits.[1] Given their large returns, low take-up of such behaviors is often seen as a puzzle that merits intervention. In recent years, both traditional interventions like "sin taxes" (O'Donoghue and Rabin, 2006; Allcott et al., 2019; Farhi and Gabaix, 2020), as well as non-traditional instruments like "nudges" (Thaler and Sunstein, 2009), have been widely used. Many have been shown to be effective and cost-effective with respect to their target outcomes.[2]

Most evaluations measure the behavior targeted by the intervention and little else unless there are specific reasons to expect spillovers.[3] But there may be general reasons to worry about spillovers. In particular, attention and effort may be costly or limited in ways that affect how beneficiaries respond to interventions. Several recent empirical papers have discovered unintended side-effects of interventions that appear to be consistent with this hypothesis (e.g. Medina (2020); Hall and Madsen (2020); Hussam and Oh (2021), and recent theoretical papers incorporate limits to attention (Persson, 2018; Nafziger, 2020) or preferences to maintain one's self-image (Dolan and Galizzi, 2015) in ways that predict intervention spillovers.

The contribution of this paper is threefold. First, I provide a simple framework for understanding how interventions that target one behavior can generate spillovers to other behaviors, whether they are rooted in limited attention, preferences, or both. I do not focus on the distinction between limited attention and preferences, but offer an alternative classification that captures key policy-relevant features of spillovers. Second, I test for the existence of spillovers in a strategically chosen context: interventions that encourage daily changes, involving minimal time, in health and wellness behavior. The context is (inevitably) specific, but one where the use of nudges is widespread, and one where traditional economic

[1] For example, see Afshin et al. (2019), Institute for Health Metrics and Evaluation (2018), and Zhou et al. (2014).

[2] For example, soda taxes have shown to reduce sugar consumption due to soda by 18% on average, and by 40% among young people age 13-21 (Dubois et al., 2020). Prompting people to write down a plan for getting the flu vaccine raised immunization rates by 13% (Milkman et al., 2011). Other interventions have proven less effective: encouraging the use of clean cookstoves has proved to be much more difficult than expected (Hanna et al., 2016).

[3] See for example the literature how retirement savings policies affect other financial behavior (e.g. Chetty et al. (2014); Besheárs et al. (2021)) and the literature on the general equilibrium effects of cash transfers and other anti-poverty programs (e.g. Egger et al. (2019); Muralidharan et al. (2017)).

models would not predict spillovers. I implement a large, pre-registered field experiment and collect objective, high-frequency data using smartphone apps. Because the experiment design involves the promotion and measurement of two behaviors, I can test for interactions as well as assess the costliness of spillovers from the perspective of a social planner who cares about both behaviors. Third, and finally, I decompose spillovers using the classification I proposed, and consider the implications for policy design.

The goal of my framework is to highlight policy-relevant features of behavioral spillovers and to generate predictions to motivate an experiment. A decision-maker (DM) has two behaviors available to her, x and y. Doing each behavior generates a return, but requires mental effort, which is costly. An outside actor can subsidize effort to one or both behaviors with incentives or SMS messages. I model three types of mechanisms that can potentially drive spillovers, which I call "overload," "diversion," and "depletion." Overload can be summarized as "intervention y affects intervention x;" diversion can be summarized as "intervention y affects behavior x," and depletion can be summarized as "behavior y affects behavior x." Cross-disciplinary evidence supports the plausibility of each spillover type. Importantly, in my model depletion will be not be separately identifiable from (positive or negative) complementarity between behaviors in the utility function. In other words, I will not aim to distinguish spillovers rooted in limited attention from those rooted in preferences regarding one's self-image.

The model generates three predictions relating these mechanisms to comparative statics of behavior with respect to messages and incentives. This motivates an experiment design with five treatment groups: a control group, a group that gets messages about behavior x, a group that gets messages about behavior y, a group that gets both sets of messages, and a group that gets incentives for behavior y. I run an online experiment, recruiting 3,845 individuals via Facebook Ads that promote a study about daily meditation (behavior x) and meal logging (behavior y). Participants took a baseline survey and downloaded two smartphone applications, one for tracking meditation, and the other for logging meals. Upon verifying that participants downloaded the apps, they were enrolled and randomized. The treatment period lasted four weeks. Participants in groups with only x or y messages received prompts twice a day targeting the corresponding behavior. Participants in the group with both x and y messages received the union of both message sets, resulting in four daily messages. Lastly, participants in the y incentive group received an expected reward for every day they successfully logged their meals. I continued to measure behavior via the apps for an additional four weeks after the end of treatment.[4]

The reduced form results show large target effects of both message and incentive interventions. Meditation messages raised the rate of meditation by 8.8 percentage points (almost double the control rate) and nutrition messages raised the rates of meal logging by 16.6 percentage points (more than double the control rate). Incentives for meal logging had an even larger effect, raising rates of meal logging by 38.1 percentage points (more than triple the control rate). But all three treatments imposed substantial spillovers on the opposite behavior, as measured by comparisons with the control group. Messages about meditation reduced meal logging by 2.4 percentage points (19%), messages about nutrition reduced meditation by 2.8 percentage points (29%), and incentives for meal logging reduced meditation by 2.5 percentage points (27%). The group with both target and non-target messages also did worse on target behaviors than the group that received just target messages, by 2.2 and 5.0 percentage points for meditation and meal logging, respectively.

To what extent do overload, diversion, and depletion explain the observed spillovers? There was no evidence of an interaction effect between the two sets of messages, which, according to the model, indicates that there is no evidence of overload with respect to behavior. But I have three additional sources of data that shed light on this mechanism: opt-out behavior, reading of messages (as measured by a surprise raffle), and recall of messages (as measured by an information quiz). I do not find strong evidence that participants with two sets of messages opt-out more than those with just one, but I do find evidence that messages generate negative spillovers on the reading and recall of non-target messages. This result is suggestive of overload on the level of information acquisition.

As is made clear by the predictions of the model, the fact that all three interventions generated negative spillovers on the opposite behavior provides strong evidence of either diversion, depletion, or both. Moreover, the fact that all three interventions generated different target effects but similarly sized spillovers sug-

[4]This is not technically a natural field experiment in the framework of Harrison and List (2004), as subjects filled out a consent form and knew they were taking part in a study. They also knew that assignment to programs was random, for reasons that I explain in Section 7.2. However, it is close to a natural field experiment since the subjects undertake the tasks in their natural environments (and the commodity, task, and information set all have a field context).

gests that depletion alone cannot drive the results. If it did, the more effective interventions should have evoked more costly effort and generated larger spillovers. Consistent with this, when I parameterize and estimate the model, I find, with 95% confidence, that diversion is a necessary component of the observed message spillovers. The key upshot is that interventions that generate larger effects on target behaviors will not necessarily generate comparatively larger spillovers. In other words, the spillovers I observe act, in part, as a fixed cost. This means that for a policy maker evaluating interventions based on classic cost-effectiveness analysis, not only are interventions likely to be more costly than they appear, but there may also be unappreciated differences in cost-effectiveness across interventions.

This paper contributes to several strands of literature. The first is the small but growing literature on behavioral spillovers rooted in limited attention, an idea recently formalized in Nafziger (2020). Motivated in part by the vast empirical research on the effectiveness of using marketing-inspired tactics to capture attention and draw it to important economic choices (e.g. Karlan et al. (2016); Taubinsky (2013); Allcott and Rogers (2014); Rogers and Milkman (2016)), researchers have begun to document that such campaigns are also capable of diverting attention. Medina (2020) found that sending SMS reminders to bank clients effectively reduced late fees paid by 14%, but it also increases overdraft fees paid by 9%, resulting in a net loss for some. Hall and Madsen (2020) found that highway safety campaigns displaying roadside fatality counts is so effective at seizing attention that it actually increases the number of traffic crashes. Altmann et al. (2019) found that in the lab, choice-promoting interventions improved performance on a target task, but it came at the expense of performance on a background task. These studies fall within the broader, rich literature on the implications of limited attention for economics (e.g. Sims (2003); Chetty et al. (2009); Bordalo et al. (2012); Gabaix (2014); Taubinsky and Rees-Jones (2018); Farhi and Gabaix (2020); Bronchetti et al. (2020); Morrison and Taubinsky (2019))

Second, this paper contributes to the more established literature on behavioral spillovers rooted in preferences. Dolan and Galizzi (2015) provide a useful framework, organizing spillovers by whether they are positive and "promote" other good behaviors, or whether they are negative and "permit" negligence.[5] Spillovers may be positive if good behavior provides a "positive signal" for one's future self

[5]In the case of a bad initial behavior, spillovers can be similarly be negative and "precipitate" other bad behaviors, or positive and "purge" them.

(Bénabou and Tirole, 2004), building a sense of identity around the behavior (Galizzi and Whitmarsh, 2019). On the other hand, if people aim to simply maintain a positive self image, then engaging in something "good" can potentially license one to subsequently engage in something "bad," or vice versa, a phenomenon that has been termed moral licensing. Empirically, some studies find interventions promoting good behaviors to have negative spillovers on other behaviors (Wertenbroch, 1998; Khan and Dhar, 2006; Tiefenbeck et al., 2013; Dolan and Galizzi, 2014) and even on the same behavior in a different context (Hussam and Oh, 2021). But other studies find positive spillovers (see for example Jessoe et al. (2017); Ek and Miliute-Plepiene (2018); Brandon et al. (2019)).[6] A related literature looks at whether fundraising appeals for one charity have negative spillovers on donations to others; most studies do not find strong negative spillovers (Scharf et al., 2017; Donkers et al., 2017; Meer, 2017; Deryugina and Marx, 2021).

Finally, this paper contributes to a growing literature that attempts to explore unanticipated consequences of nudges, both in general equilibrium (Spiegler, 2015) and for welfare more broadly. Several recent papers have taken more seriously the psychological costs of nudges that exploit social comparisons, shaming, or social pressure (Allcott and Kessler, 2019; Butera et al., 2019; DellaVigna et al., 2012; Jimenez-Gomez, 2018), in order to measure total welfare effects. Another recent literature has documented substantial opt-out behavior in text message programs, indicating that messages can be costly (Damgaard and Gravert, 2018; Fricke et al., 2018). Nudges have also been shown to exacerbate adverse selection in health insurance markets (Handel, 2013), impose hidden time costs (Taylor, 2020), and weaken support for costlier policies by offering a promise of a "quick fix" (Hagmann et al., 2019). Most of these studies do not dispute the benefits of nudges, but rather aim to expose their costs for the sake of a fuller picture.

The present study builds on the above literature in two ways. First, it tests for spillovers in a field context more natural than Altmann et al. (2019), but more generic than existing natural experiments, with sufficient control to randomly promote both of the measured behaviors within the sample.[7] This provides a more complete picture of spillovers, including the testing of interaction effects as well as

[6] Galizzi and Whitmarsh (2019) provide a useful review of the empirical methods used to measure spillovers.

[7] For example, in Medina (2020), the observed spillovers can be attributed to the budget constraint. The findings in Hall and Madsen (2020) are quite specific to billboards and traffic accidents. Neither paper randomly both promotes and measures two behaviors.

an assessment of spillover costs. Second, this study exploits a framework and experiment design that help illuminate the types of spillovers at play, and how they might interact with cost-effective policy design.

I begin in Section 2 by describing the model and its predictions for behavior with respect to incentives and messages. In Section 3 I describe the experiment. In Section 4 I present orthogonality tests, descriptive statistics, and reduced form results on behavior, expectations, opting out, and the reading of messages. In Section 5 I describe and estimate the structural model, use my parameter estimates to decompose spillovers. In Section 6 I discuss implications for cost-effectiveness analysis. In Section 7 I discuss alternative explanations for the observed spillovers, including time constraints, the bundling of treatments, and experimenter demand effects, and in Section 8 I conclude.

2 Framework

I use a simple framework to derive comparative statics of behavior with respect to messages and incentives. In this section I introduce the model, discuss the relevant evidence for costly effort and attention, and generate three predictions that will be testable with the reduced form results of the experiment.

2.1 Set-Up

I consider an agent who chooses how much cognitive effort a_j to invest in two behaviors, x and y.[8] She receives return u_j per unit of effort invested in action $j \in \{x, y\}$. She also receives u_{xy} per joint unit invested, allowing for (positive or negative) complementarity between behaviors x and y; this can also capture the idea of "moral licensing" described in Section 1. The agent faces some effort cost, which I denote by the function C; this captures all cognitive effort costs, including those to both attend to and engage in the activity. Let the agent's utility over effort be $U(a_x, a_y) = a_x u_x + a_y u_y + a_x a_y u_{xy} - C(a_x, a_y)$, her returns from exerted effort minus its costs. She chooses a_x and a_y to maximize her utility.

[8] I use "a" instead of "e" to emphasize that this is cognitive effort, which can also be understood as attention according to some definitions. For example, Chun et al. (2011) define attention in a way that includes everything from information processing to executive function to self-control. Throughout the paper I will use the terms "attention" and "cognitive effort" interchangeably.

$$\max_{a_x, a_y} \left\{ a_x u_x + a_y u_y + a_x a_y u_{xy} - C(a_x, a_y) \right\} \tag{1}$$

Effort is a latent variable; I observe only whether the agent does actions x and y in each period. I will assume that due to random error, I observe the agent undertaking action j as long as $a_j + \xi_j > 0$, where ξ_x and ξ_y are i.i.d. shocks that are independent from one another and conditionally uniformly distributed. The probability of observing, for example, $x = 1$ is thus $Pr(a_x + \xi_x > 0)$.[9]

Effort costs cause the agent to expend less effort than she otherwise would. However, an outside actor can introduce an intervention w_j for behavior j, where the intervention can be either messages m or incentives z ($w_j \in \{m_j, z_j\}$).[10][11] Let the modified cost function be $C(a_x, a_y, w_x, w_y)$. Let the marginal cost of a_x, C_1, be denoted as $c^x(a_x, a_y, w_x, w_y)$ and the marginal cost of a_y, C_2, be denoted as $c^y(a_y, a_x, w_y, w_x)$. I make the following assumptions about the cost function.[12] First, I assume that $c^x_1 > 0$ and $c^y_1 > 0$.[13] Second, I assume that both target messages and target incentives act as effort subsidies for the target behavior, reducing the marginal cost of effort: $c^x_3 < 0$, $c^y_3 < 0$.[14]

In what follows, I define three mechanisms that will be capable of driving spillovers: overload, diversion, and depletion.

Definition 1. ***Overload*** *is the event that* $c^x_{34} > 0$ *or* $c^y_{34} > 0$

[9] Note that if the returns to a behavior are non-positive, the agent will simply allocate zero effort and will abstain from the action. In this sense, the model is in the spirit of the rational inattention literature: the agent only pays attention if she has something to gain (though here the costly attention or effort is paid to the behavior itself, rather than to an information structure). I do assume, however, that agents must expend effort or attention in order to do the behavior; this is consistent with these being behaviors that are not currently habitual, as will be the case for participants in the experiment.

[10] I assume continuous treatments m and z here, but in the empirical and structural models, treatments will be binary.

[11] In this framework, there is no specific market failure, but by assuming that a policymaker chooses to intervene, I am implicitly assuming one. For example, if time inconsistent agents continually postpone the high effort costs of initiating a healthier diet, a policymaker can potentially solve this problem by subsidizing effort costs in the present. The focus of this paper is not market failures, but the potential unintended consequences of interventions that might solve them.

[12] Subscripts denote the derivative with respect to the corresponding argument; note that the order of arguments in c^x and c^y are different, so that, for example, c^j_1 always represents the derivative of the marginal cost of a_j with respect to behavior j.

[13] This amounts to the standard assumption that effort costs to individual behaviors are convex, and also ensures the existence of a local maximum.

[14] I assume that the outside actor will not implement both messages and incentives. The framework does not have interesting implications for interactions between messages and incentives, so I do not implement this treatment in my experiment.

In the presence of overload, y messages interfere with the subsidy produced by x messages (and vice versa).

Definition 2. ***Diversion*** *is the event that* $c_4^x > 0$ *or* $c_4^y > 0$*.*

Diversion thus operates like a tax: messages or incentives about behavior y increase the marginal cost of effort to x (and vice versa). I allow for the possibility of both message and incentive diversion.

Definition 3. ***Depletion*** *is the event that* $c_2^x = c_2^y > 0$*.*[15]

Depletion implies that the marginal cost of exerting effort to do behavior x is increasing in the effort expended on behavior y, and vice versa.[16]

Overload, diversion, and depletion essentially capture different functional forms that behavioral spillovers might take. Overload can be summarized as "intervention y affects intervention x;" diversion can be summarized as "intervention y affects behavior x," and depletion can be summarized as "behavior y affects behavior x." These are new concepts that have been constructed for the specific purpose of distinguishing between spillover types with distinct policy implications. That being said, there is a wealth of evidence on attention and effort across several disciplines that support the plausibility of each spillover type.

2.2 Overload, Diversion, and Depletion in the Literature

In this section, I discuss theory and evidence, from psychology and other disciplines, that underpin the ideas of overload, diversion, and depletion. I employ a taxonomy of attention that distinguishes between external versus internal attention (Chun et al., 2011).[17] External attention refers to the selection and modulation of external information, and the storing of that information in the brain.[18] Internal

[15]The equality is due to the fact that $c_2^x = \frac{\partial^2 C}{\partial a_x \partial a_y} = \frac{\partial^2 C}{\partial a_y \partial a_x} = c_2^y$. For this reason, from now on I will write $c_2^x = c_2^y$ as simply c_2^x.

[16]I do not rule out the possibility of positive spillovers. In the self-signaling models (e.g. Bénabou and Tirole (2004)), for example, depletion would be hypothesized to be positive, and indeed I ultimately obtain positive (though highly uncertain) estimates of depletion. However, given that I hypothesized negative spillovers in my pre-analysis plan (see Appendix J), and that this is what I ultimately find, I focus on mechanisms that could explain negative spillovers.

[17]The roots of this taxonomy go back to 1890, when William James distinguished between "passive" and "active" attention. This basic division has persisted over the years, with several variants—bottom-up versus top-down attention, stimulus-driven versus goal-driven attention, exogenous versus endogenous attention, and finally external versus internal, which is what I use here.

[18]It can be directed to the sensory modalities of sight, hearing, touch, smell, and taste, and it can be used to perceive the world across space ("spatial attention") or time ("temporal attention").

attention, on the other hand, refers to the selection and modulation of content that has already been stored in the brain.[19] I build on this taxonomy to describe the possible ways in which interventions that require little time and no physical effort could impose negative spillovers, summarized in Appendix Figure A1. Suppose there is an intervention, say a text message, about some behavior x. External attention is used to modulate that intervention and store it in our brains. Internal attention is then used to think about x, and ultimately do x. Now, suppose there is also a text message about some different behavior y. Three things might happen. First, limits to external attention, or limited information processing, might cause the y intervention to interfere with the x intervention. This is what I have defined to be "overload." Second, limits to working or short-term memory might cause the y intervention to divert attention toward y and away from x, reducing the likelihood of doing x (regardless of whether or not there is any x stimulus). This is what I have defined to be "diversion." Finally, if the y intervention works, causing us to do y and to exert costly cognitive effort or time, we might be subsequently less likely to do x. This is what I have defined to be "depletion."

There is ample research on overload across several disciplines. The fact that people are limited in their ability to process stimuli is so well-established that words like "selection" are commonplace in the psychology literature; the relevant question is not whether we select which stimuli to process but how.[20] In the economic arena, we have evidence that people have difficulty processing all of the information about the products they buy (Lacetera et al., 2012), all of the choices available to them (Chernev et al., 2015), and all of the dimensions of their production processes (Hanna et al., 2014).[21] With respect to messaging, there is evidence that people become habituated or desensitized to alerts, designed to promote some behavior, over time.[22] The possibility of diversion is also well supported in the lit-

[19] Internal attention includes the attention required to think about, plan, and make decisions about an action—including executive function, working memory, and long-term memory. It also includes the cognitive effort and self-control required to carry out a task.

[20] See for example one seminal paper which shows that limits to external attention are modality-specific: people are unable to attend to two visual or two auditory streams (Neisser and Becklen, 1975), but better able to attend to one of each (Duncan et al., 1997).

[21] In particular, the "information overload" literature in marketing has documented a hump-shaped relationship between the quantity of information that consumers have about products, and the "decision quality" of their ultimate purchase (Hwang and Lin, 1999; Edmunds and Morris, 2000; Eppler and Mengis, 2004).

[22] For example, in medicine, as the use of electronic medical records and attendant automatic alerts to provide "decision support" have become widespread, there has been extensive discussion of "alert fatigue," the idea that physicians become habituated to alerts over time. One SMS program designed

erature. Experiments on working memory show that focusing on one thing often comes at the expense of something else.[23] Studies on "attentional capture" show that irrelevant stimuli can easily draw people's attention away from a task at hand (Yantis and Jonides, 1984).[24,25] With regard to depletion, the psychology literature has recently made headway in demonstrating the idea of costly cognitive effort. Botvinick and Braver (2015) found that performance on difficult tasks tends to increase with incentives, and Dunn et al. (2016) showed that when given a choice between tasks that require high and low cognitive effort, participants tend to prefer the latter.[26]

2.3 Predictions

I now derive comparative statics of effort levels a_x^* and a_y^* with respect to messages and incentives, which will be randomly varied in the experiment. Since I have assumed that ξ_x and ξ_y are conditionally uniformly distributed, I can estimate these comparative statics with linear probability models of behavior on treatments.[27]

Importantly, I will be unable to separately identify $c_2^x = c_2^y$ from u_{xy}, meaning that I will not be able to distinguish between (positive or negative) complementarities in the utility function and depletion in the effort cost function. Since moral licensing would be captured by a negative complementarity in the utility function, this means that I cannot separately identify depletion from moral licensing. For this reason, in the remainder of the paper, I will drop the u_{xy} term, and simply note that whenever I refer to depletion, I am really referring to depletion net of any complementarity in the utility function. This will have important implications for the interpretation of my results, which I will discuss in Sections 5 and 6.

I begin by defining two important objects: negative spillovers and interference.

to alert physicians to new clinical trials found that response rates declined 2.7% every two weeks (Embi and Leonard, 2012).

[23] For example, people are capable of holding only limited sets of digits or words at a time (Miller, 1956; Luck and Vogel, 1997).

[24] The types of stimuli most likely to achieve attentional capture are novel stimuli (i.e. an unexpected SMS), emotionally salient stimuli (i.e. footage of a humanitarian crisis) and stimuli associated with rewards (i.e. a plate of cookies placed in front of you) (Fawcett et al., 2015; Chun et al., 2011).

[25] One relevant example is the phenomenon of "intention cost," or reduced performance (and brain activity) in a current task as a result of thinking ahead to a future task (Burgess et al., 2003; Gonen-Yaacovi and Burgess, 2012).

[26] Typically, the more automatic the task—the closer it is to some "default" behavior—the less effort it requires (Shenhav et al., 2017). Given that the policy goal of "behavior change" inherently asks people to move away from their defaults, it may also inherently require cognitive effort.

[27] Probit and logit models give similar results.

Definition 4. *An intervention promoting behavior j generates a* ***negative spillover*** *on behavior i when $\frac{\partial a_i^*}{\partial w_j} < 0$.*

In other words, a negative spillover is a negative response of a behavior to a non-target intervention. If $\frac{\partial a_x^*}{\partial m_y} < 0$, it constitutes a *message spillover*, and if $\frac{\partial a_x^*}{\partial z_y} < 0$, it constitutes an *incentive spillover*.

Definition 5. *Two interventions promoting behaviors i and j generate* ***interference*** *when $\frac{\partial^2 a_i^*}{\partial m_i \partial m_j} < 0$.*

In other words, interference is a negative interaction between two interventions. If $\frac{\partial^2 a_x^*}{\partial m_x \partial m_y} < 0$, it constitutes *interference*, and if both $\frac{\partial^2 a_x^*}{\partial m_x \partial m_y} < 0$ and $\frac{\partial^2 a_y^*}{\partial m_x \partial m_y} < 0$, it constitutes *interference in both directions*.

I obtain three key predictions. The first relates spillovers to the presence of depletion and/or diversion.

Proposition 1. *Either message (incentive) diversion or depletion is a necessary condition for message (incentive) spillovers, and the presence of both message (incentive) diversion and depletion is a sufficient condition for message (incentive) spillovers.*

Proof. The proof is straightforward from the expression for $\frac{\partial a_x^*}{\partial w_y}$, Equation 8 in Appendix B.1. □

Proposition 1 implies that if I find message or incentive spillovers, they must be due either to diversion or depletion. If I find neither, then the conclusion is ambiguous.

The second prediction will help distinguish between diversion and depletion.

Proposition 2. *If both message and incentive diversion are absent but depletion is present, then:*

$$\frac{\partial a_i^*/\partial m_j^*}{\partial a_j^*/\partial m_j^*} = \frac{\partial a_i^*/\partial z_j^*}{\partial a_j^*/\partial z_j^*} \tag{2}$$

Proof. See Appendix B.2. □

In words, if diversion is absent, then the ratio between spillover and target effects generated by an intervention is fixed across the two intervention types: messages and incentives. This means that if we can reject equality of these ratios, we can be confident that spillovers are driven at least in part by diversion. The intuition for this proposition is the following. Recall that I have defined depletion to be the possibility that cognitive resources (effort, self-control, executive function) required to take action are costly. Since a spillover driven by depletion operates through the target action, we should expect interventions with large positive target effects to also have large negative spillovers, and interventions with small positive target effects to have small negative spillovers. These spillover/target effect ratios should be constant across different interventions. (If both diversion and depletion are absent, then the spillover will still be constant across intervention types at zero.)

The third prediction relates interference to overload.

Proposition 3. *Assume that $c^x_{34} = c^y_{34}$ and that all other second derivatives of c^x and c^y are zero. Then overload is a necessary condition for interference in both directions, and a sufficient condition for interference in one direction.*

Proof. See Appendix B.3. □

The implication of Proposition 3 is as follows. If I find interference in both directions, it must be due to overload. If I find interference in neither direction, then there is no evidence of overload. And if I find interference in just one direction, the conclusion is ambiguous (and the second assumption does not hold).

As indicated, Proposition 3 requires two additional assumptions. First, I assume that, with the exceptions of c^x_{34} and c^y_{34}, the second derivatives of c^x and c^y (third derivatives of C) are zero.[28] Second, I assume that $c^x_{34} = c^y_{34}$; namely, that the overload effect is symmetric across different behaviors. This implies that if two messages are sent, one about x and one about y, the first will interfere with subsidy generated by the second on y just as much as the second interferes with the subsidy generated by the first on x. Ultimately I can check this assumption by testing whether $\frac{\partial^2 a^*_x}{\partial m_x \partial m_y} = \frac{\partial^2 a^*_y}{\partial m_x \partial m_y}$. This test is not rejected at the 5% level, but the p-value is 0.09. I will take this into account when I discuss the outcome of this test, and what we can learn from it, in Section 4.

[28] I have no reason to believe that these derivatives are first order, as economic intuition tells us nothing about them.

3 Experiment Design

3.1 Design and Protocol

The experiment design is displayed in Table 1.[29] The control group received no intervention. Group 2 received only messages about behavior x, and Group 3 received only messages about behavior y. Group 4 received messages about behavior x as well as messages about behavior y. Group 5 received incentives for behavior y. Behaviors x and y were daily meditation and nutritional self-monitoring via meal logging. These behaviors were chosen for three reasons. First, they are important health behaviors for the sample frame (young Americans).[30,31] Second, both behaviors can be measured objectively at high frequency via pre-existing smartphone applications.[32] Third, both behaviors require minimal amounts of time. The average meditation session was 21 minutes, but meal logging only took 11 minutes per day on average. Although this does not rule out the time constraint as a driver of spillovers, it does make it less likely. I will address this possibility in Section 7.1.

Participants were recruited using Facebook advertisements, targeting adults age 18-35 living in the U.S. (see Appendix C Figure A2). Upon clicking the link, participants underwent a brief screening that ensured they (1) had an iPhone or android; (2) were over 18; (3) were interested in working on wellness habits like meditation and tracking nutrition; and (4) were willing to download the two free applications. They then provided informed consent and proceeded to Survey 1, which took about 15 minutes. The first part of Survey 1 provided instructions for downloading the two apps (which were also emailed upon survey completion). Participants were instructed that in order to enroll, they would need to download both apps within 24 hours. Participants were then asked questions on demograph-

[29]The study was registered at the AEA RCT registry with a pre-analysis plan. Please see Appendix J for details.

[30]A recent meta-analysis in the Journal of the American Medical Association found that meditation programs improved anxiety by 0.38 SDs at 8 weeks (and 0.22 at 3-6 months), improved depression by 0.30 SDs at 8 weeks (and 0.23 at 3-6 months), and reduced pain by 0.33 SDs at 8 weeks (Goyal et al., 2014). The use of smartphone apps for nutritional self-monitoring and feedback have been linked to weight loss (Wharton et al., 2014), which is associated with many health benefits.

[31]In the screening process, I asked participants if they would feel comfortable using a nutrition tracking app, and asked those who have struggled with eating disorders or body image issues in the past to consider this carefully.

[32]The meditation application allowed participants to access a wide variety of guided meditations or meditate on their own, and recorded details about each meditation session including the time spent meditating. In the meal logging application, participants inputted information about the meals they ate and then tracked various measures of the nutritional quality of their diet.

ics, electronic notifications, and preferences/experiences surrounding meditation and meal logging.

Participants who were verified to have downloaded both apps (3,855 of the 5,845 who filled out Survey 1, or 66%) were then randomized to treatments, re-randomizing on gender, age, whether or not they had a college degree, daily notifications, whether or not they meditated in the last month, and whether or not they tracked their meals in the last month. These participants then received an enrollment confirmation email with their treatment assignment, a link to Survey 2, and other details about the study. Survey 2 required about five minutes and contained questions about participants' expectations of each behavior, conditional on their treatment assignment.

Importantly, when informed of their treatment assignment, participants were told that "this assignment was completely random, and has nothing to do with your survey responses or the relative importance of meditation, exercise, nutrition, and sleep." The purpose of this was to rule out an alternative potential source of spillovers or interference: the possibility that participants infer the relative benefits of the behaviors from their treatment assignment.[33] In order to avoid experimenter demand effects, I also tell participants that, "Depending on your above assignment, we may (or may not) be encouraging you to meditate and/or log your meals, but your ultimate use of the apps is entirely up to you." Finally, and relatedly, I give the meditation and meal logging programs separate names (Remindful and eNOMerate, respectively), and send the messages from different phone numbers, in order to "unbundle" the two behaviors as much as possible. I discuss these concerns further in Section 7.2.

Each message program included twice-daily text messages: one simple reminder to do the behavior, and one longer message with information about some proven benefits to the behavior, as demonstrated in Table A1. As the table demonstrates, the timing of meditation vs. nutrition messages and information vs. reminder messages alternated in a balanced fashion. Messages were sent at either 7am and 7pm or at 8am and 8pm, alternating on a daily basis, and scheduled so that the group that received both meditation and nutrition messages never received them at exactly the same time (it was always the case that one message was at 7 and the other at 8). The purpose of this was to avoid capturing mechanical interference due

[33] This mechanism is potentially important, but cannot be well studied in this kind of experimental context, since many participants already assume assignment is random.

to the simultaneous arrival of messages. There were 14 distinct messages, and 27 days of treatment, so each message (save one) was sent twice over the course of the program. The full set of messages is shown in Appendix Table A2. Participants were told during consent, in the enrollment email, and at the start of treatment that they could opt out any time by replying "STOP" to the relevant number.

Incentives took the form of a raffle. Participants were informed in their enrollment email that they would earn one green lottery ticket for every day they successfully do the behavior, and one red lottery ticket every day that they do not. We informed them that at the end of the four weeks, we would draw one of their tickets, and a green ticket would be worth a $10 Amazon gift certificate. Every Sunday during the program, participants received an email updating them about the tickets earned the previous week. They also received an email informing them when the program ended.

Four weeks after the end of treatment, participants received Survey 3 via email. Survey 3 included questions about meditation and meal logging outside of the assigned apps, the timing of behaviors, some measures of mental health and diet, and quizzes about the information content of any message program they received. For further details about the experiment protocol see Appendix C. Of the 3,855 who were enrolled in the study, 0.3% withdrew from the study, resulting in a final sample of 3,845, For further details about attrition, see Appendix D. Of the 1,585 assigned to meditation messages, 13% opted out, and of the 1,625 assigned to meal logging messages, 18% opted out. I analyze this data in Section 4.4.

The sample used for the experiment is obviously very particular and select, dominated by relatively young college educated women with prior experience engaging in these behaviors. This was by design: I recruited people who were likely to be compliers in order to raise the likelihood of having strong target treatment effects, a critical precondition for studying spillovers and interactions. I discuss external validity concerns more generally in Section 7.3.

3.2 Data

I have four sources of data. The first is data on meditation and meal logging from the two apps. The meditation app records minutes meditated on the app, and the meal logging app provides data on whether at least one meal was logged as well as total calories logged.[34] For much of the analysis, I focus on binary outcomes—whether or not individuals meditated each day, and whether or not they logged at least one meal—because this is what the treatments promoted. Messages focused on daily behavior at the extensive margin, sometimes even downplaying the importance of intensity (i.e., "find a few minutes to meditate / log your meals"). Incentives were explicitly tied to the extensive margin, as described above. Correspondingly, these binary indicators were the principal outcomes specified in the pre-analysis plan.

There are a few limitations to the data collected from the apps. First, with the meditation app, it is possible to record meditation without actually meditating (the person can start the timer and do something else). For this reason, I did not include incentives for meditation as a treatment. Still, it is possible that some part of the meditation message treatment effect is driven by participants induced to falsely report meditation as a result of messages. To the extent that this happened, I would be measuring both effort to meditate, and effort to falsify meditation, presumably motivated by experimenter demand effects. Second, neither the meditation nor meal logging measures capture activity done outside of the apps.[35] Finally, the measure of calories logged is not an unbiased measure of meal logging, since it conflates meal logging and dietary changes, which further supports my focus on binary outcomes.

The second source of data comes from the three surveys described above. The third is individual-level data on opt-out for those assigned to messaging treatments. The fourth is data on participation in a surprise raffle that I conducted with those assigned to messaging treatments, described in Section 4.5, which I use as a proxy for message reading.

Table 2 shows means and standard deviations of key variables across treatments, as well as an F-test of the joint significance all treatment variables. The re-randomization procedure ensured that the first variables were balanced across treatments, and the rest of the variables are highly balanced as well. Overall, the sample was overwhelmingly female (93%), mostly college educated (71%), with an average age of 27. Participants receive on average 51 notifications daily, 36 of which are messages, 10 of which are updates, 4 of which are reminders, and 1 of

[34] Due to technical constraints, I do not have data on the number of meals or items logged or what foods were logged.

[35] In Survey 3 I asked participants how many days they did each behavior outside of the apps; in Table A8 I inflate meditation and meal logging rates accordingly and find similar results.

which was classified as "other." (See Appendix H, Figure A6 for details.) Most participants had experience with both meditation and meal logging. With respect to meditation, 90% had meditated before, 57% had done so on a daily basis, and 46% had done so in the last month. With respect to meal logging, 90% had logged their meals before, 87% had done so on a daily basis, and 32% had done so in the last month. (See Appendix H, Figure A7 for details.) These are people who have strong prior interest and experience in both behaviors, but who, for whatever reason, have not been engaged in them recently. In Section 6.3, I discuss implications of the self-selected nature of the sample.

4 Reduced Form Results

For my main specification, I estimate linear probability models of the outcome on treatments at the individual-day level, where the outcome is 1 if the participant did the behavior on a given day. For meal logging, the behavior is having logged at least one meal and 0 otherwise. I define m_x (m_y) to be 1 if the individual received x (y) messages and 0 otherwise; m_x*m_y is 1 if the individual received both sets of messages.

The main goal here will be to estimate spillovers and interference in order to test Propositions 1-3. Since in the experiment, treatments are binary, I redefine spillovers and interference as follows. An intervention w_j promoting behavior j generates a negative spillover on behavior i when $\mathbb{E}[a_i|w_i = 0, w_j = 1] - \mathbb{E}[a_i|w_i = 0, w_j = 0] < 0$. Two messaging interventions promoting behaviors i and j generate interference when $(\mathbb{E}[a_i|m_i = 1, m_j = 1] - \mathbb{E}[a_j|m_i = 0, m_j = 1]) - (\mathbb{E}[a_j|m_i = 1, m_j = 0] - \mathbb{E}[a_j|m_i = 0, m_j = 0]) < 0$.[36] I estimate spillovers and interference using the below specifications (where i indexes the individual and t indexes the day):

$$x_{it} = \beta_0^x + \beta_1^x m_x + \beta_2^x m_y + \beta_3^x m_x m_y + \beta_4^x z_y + Z_i + \psi_t + \epsilon_{it} \tag{3}$$

$$y_{it} = \beta_0^y + \beta_1^y m_y + \beta_2^y m_x + \beta_3^y m_x m_y + \beta_4^y z_y + Z_i + \psi_t + \epsilon_{it} \tag{4}$$

$\hat{\beta}_2^x$ is an estimate of the spillover of m_y on x; $\hat{\beta}_2^y$ is an estimate of the spillover of

[36] In the parameterized version of the model in Section 5.1, the two sets of definitions are equivalent. In other words, spillovers and interference defined as partial derivatives with respect to continuous treatments, as in Section 2, are equal to the differences in conditional means shown here.

m_x on y, and $\hat{\beta}_4^x$ is an estimate of the spillover of z_y on x. $\hat{\beta}_3^x$ and $\hat{\beta}_3^y$ are estimates of interference. I include a vector of controls Z_i that consists of the variables used for re- randomization: whether or not the participant is female, whether or not they completed college, daily notifications, whether or not they meditated in the month prior to the study, and whether or not they logged a meal in the month prior to the study. I also include day fixed effects ψ_t, and cluster standard errors at the individual level. Coefficients on message treatments represent intent-to-treat effects, as some participants chose to stop receiving messages.

4.1 Main Effects

I begin by discussing binary outcomes during the treatment period. The raw data is shown in Figure 1 and the regression results are shown in columns 1 and 2 of Table 3. Recall that I report main effects of m_x, m_y, and z_y as well as an interaction term between m_x and m_y. (See Appendix H Table A7 for estimates reported as mutually exclusive treatment effects.) In the middle panel of the table I report linear combinations that capture the difference between being assigned both messaging treatments versus just one, and at the bottom of the table I report control means.

All three treatments—m_x, m_y, and z_y—had large and significant target effects. The coefficient on m_x in column 1 shows that meditation messages raised the meditation rate from 9.4% to 18.2%, and that on m_y in column 2 shows that nutrition messages raised the meal logging rate from 11.8% to 28.4%. Incentives for meal logging had large target effects, more than quadrupling the rate of meal logging (from 11.8% to 50%) as evidenced by the coefficient on z_y in column 2.

All three treatments also had significant negative spillovers on non-target behaviors. Participants getting only nutrition messages meditated 29% less than the control group (6.6% relative to 9.4%), as demonstrated by the coefficient on m_y in column 1. Participants getting only meditation messages logged meals 19% less than the control group (9.4% relative to 11.8%), as demonstrated by the coefficient on m_x in column 2. The group with both m_x and m_y did 2.2 percentage points worse on meditation than the group with just m_x, as captured by the linear combination $m_y + m_x m_y$. Similarly, the group with both m_x and m_y did 5.0 percentage points worse on meal logging than the group with just m_y, as captured by the linear combination $m_x + m_x m_y$. The incentive treatment z_y also had negative spillover effects, reducing meditation by 27% (6.9% relative to 9.4%) as evidenced by the coefficient on z_y in column 2. Given Proposition 1, the fact that m_x, m_y, and z_y all generated negative spillovers on the non-target behaviors implies that participants are subject to either diversion, or depletion, or both.

There is no evidence of any interference—interaction effects between m_x and m_y—neither for meditation nor for meal logging. According to Proposition 2, this implies that there is no evidence of overload. However, recall that Proposition 2 relied on the assumption that overload was not asymmetric, which we only weakly reject (p=0.09). I thus do not make strong conclusions about overload as a result of this test. I will return to the question of overload with other data in Section 4.5.

I test Proposition 3 by comparing spillover/target ratios across interventions m_y and z_y. For the binary outcome (meditating or logging at least one meal), this comparison gives spillover/target ratios of 0.07 and 0.17 for y incentives and messages, respectively, and I can reject that the ratios are equal (p=0.03).

In columns 3 and 4, I estimate the same specification for two continuous measures of each behavior: minutes meditated, and daily calories logged.[37] Recall that the measure of daily calories logged confounds meal logging effort and dietary changes. However, given that I see minimal effects on diet or weight loss in Table A10, I view total calories logged as indicative mainly of meal logging effort.

In column 3 we see that in the control group, participants meditated 2.1 minutes per day; those in the m_x group meditated 1.3 minutes more (or 62%). Turning to spillovers, we see from column 1 that those in the m_y and the z_y groups both meditated about 0.9 minutes less, on average. As with the binary outcomes, I find no evidence of interference between the two treatments. The results on calories logged in column 4 are generally consistent with the binary measures. In the control group, participants logged 153 calories on average; those in the m_y and z_y groups logged 212 and 351 calories more, respectively. Meditation messages generate a negative spillover of about 36 calories logged.

In Appendix Table A9, I look exclusively at the intensive margin, restricting the sample to subjects who did the behavior in that period. Interestingly, I find that meditation messages and meal logging incentives both had *negative* effects on

[37]In Table A10 in the Appendix I show treatment effects on three additional health outcomes, including self-reported mental health, fraction of participants' weight goal achieved, and self-reported diet scores. I find no significant effects. This is not necessarily inconsistent with prior research on meditation and meal logging, which typically evaluate much more intensive programs. For example, the meta-analysis of studies on meditation looked at TOT effects of meditation programs that lasted between eight weeks and six months; here I am estimating ITT effects of encouraging people to meditate over four weeks.

their respective target outcomes. So conditional on meditating, meditation messages caused people to spend 3.6 fewer minutes, and conditional on logging at least one meal, meal logging incentives caused people to record 21% fewer calories. (Whether this latter outcome is explained by people expending minimal effort to earn the incentive, or by people actually changing their diets, cannot be inferred.) These results suggest that the target effects of both meditation messages and meal logging incentives were driven by the extensive margin. Spillover effects, however, are negative on both the extensive and intensive margins (though the spillover effect of meditation messages not significantly so).

Testing Propositions 1 and 2 using continuous outcomes result in similar conclusions. If I test Proposition 3 using continuous outcomes, I find a difference in spillover/target ratios across meal logging incentives and messages, but it is only marginally significant at the 10% level (p=0.097). This is because while meal logging incentives had more than double the effect of messages on the extensive margin of logging at least one meal, they were only 65% more effective than messages on the intensive margin of calories logged. This makes sense: the meal logging incentives were explicitly tied to logging at least one meal. Indeed we see in Table A9 that conditional on logging at least one meal per day, participants with messages (and even control participants) record more calories on average then participants with incentives.

The aforementioned tests of Proposition 3 provide the first evidence in support of diversion as a driver of spillovers. The reduced form results on the covariance between x and y provide a second clue. Since I have defined depletion to be the possibility that doing one behavior raises the marginal cost of doing the other, we should expect that in the presence of depletion (net of complementarity in the utility function), doing x should negatively co-vary with doing y. With heterogeneous agents, however, the covariance between x and y also reflects any underlying covariance in preferences about meditation and meal logging. We can separate the two by comparing the covariance between x and y across treatments. Specifically, if there is no depletion (net of complementarity), then the covariance between x and y should be the same across all treatments, since it only reflects co-varying preferences which do not change with treatments. However, in the presence of depletion (net of complementarity), the covariance should vary across treatments, treatments induce different amounts of action, resulting in different amounts of depletion.

Table 4 shows the effects of treatments on the covariance between x and y. Our static model does not specify whether the predictions about the covariance refer to the covariance of x and y across people, or within a person over time. Both are plausible: depletion can conceivably cause people who meditate to be unlikely to also log their meals; it can also conceivably cause people who meditate on one day not to log their meals on the same day. Therefore, I check both: column 1 looks at the covariance over individuals (within days) and column 2 looks at the covariance over days (within individuals). If depletion drives spillovers, we should expect both sets of messages, as well as incentives, to have negative effects on the covariance. I see no evidence of this, supporting the conclusion that depletion (net of complementarity) cannot fully explain the spillovers we observe.

So far, the differences in the ratios of spillover to target effects, as well as the fact that covariances between meditation and meal logging appear constant across treatments, support the idea that diversion is an important driver of spillovers. To make further headway on the question of mechanisms, in Section 5.1 I will introduce additional parametric assumptions that will enable a more direct decomposition. First, however, I turn to four additional pieces of data that help provide a richer picture of what happened during the experiment.

4.2 Dynamics

The dynamics portrayed in Figure 1 illuminate several things. First, we see that all treatment groups had positive rates of both meditation and meal logging at the start of the study, which decayed substantially over time. The fact that in the first week of the treatment period, the control group meditated and logged their meals at rates of 14% and 24%, respectively, suggests that the initiation of the study was motivating in and of itself. But this motivation clearly wore off over time, at rates that do not seem to differ substantially by treatment. Computing treatment-level averages over participants by day, I find daily decay rates of 2.6% and 4.7% for meditation and meal logging, respectively, and no evidence that they vary across treatment groups. In other words, the initial excitement at the start of the study wore off over time, and neither messages nor incentives were strong enough to counteract it.

In Table 5 I show estimates of treatment effects from the post-treatment period. Since decay appears to be somewhat smooth over the full study period, without sharp drops at the end of treatment (with the exception of the z_y group), we should see the post-treatment period as capturing not only the end of treatment, but also the accumulated effects of decay. Control group means confirm the decline in engagement for both behaviors. During the post-treatment period, participation in meditation for the control group drops by 43% (from 9.4% to 5.4%), and meditation length drops by 29% (from 2.1 to 1.5 minutes). Both participation in meal logging (logging at least one meal a day) and the calories logged drop by about 72% (from 11.8% to 3.3%, and 153 to 41 calories, respectively).

Though treatment effects also diminish after the end of treatment, they do persist, for both target and non-target behaviors. The effects of messages on target behaviors are 28% and 18% the size of their treatment-period magnitude for meditation and meal logging, respectively. Target effects of incentives also persisted, at 10% the size of their treatment period effect. Interestingly, the negative spillover effects of meal logging messages and incentives on meditation rates persisted at almost 100% of their treatment period effects. These post-treatment spillovers cannot be explained by either depletion or diversion alone. Namely, the fact that the target effects of meal logging messages and incentives shrunk to 18% and 10% of their treatment-period effects, but maintained spillovers of the same size, is a case against depletion in the same spirit as Proposition 3. Diversion also cannot account fully for these spillovers as in the post-treatment period there is no message or incentive to divert attention.

We thus have three important facts to explain: the constant rate of decay across treatments; the persistence of both treatment effects and spillover effects after the end of treatment, and the differential persistence of spillovers. The first two facts are consistent with patterns of action, backsliding, and persistence that have been documented in a number of settings (e.g. (Allcott and Rogers, 2014; Agarwal et al., 2013)). A model of learning or habit formation, where attention or effort also wane over time, can explain these two facts, but it can also the third fact: why spillovers persist at larger shares of their treatment-period effects than do target effects. Namely: the absolute difference in meditation rates between the m_x group and the control group falls mechanically over time, much more than does the absolute difference between the control group and the m_y group. So we would expect spillovers, relative to target effects, to persist at larger shares of their treatment period rates, which is exactly what I find. In summary, the strong spillovers that persist after the end of treatment are unlikely to be due to concurrent instances of diversion or depletion. They are more likely to reflect treatment-period spillovers that linger due to learning or habit formation.

4.3 Expectations

During Survey 2, immediately after treatment assignment, but before treatment began, I collected data on the expectations of participants. Table 6 shows parameter estimates from regressions of individual-level expectations about rates of meditation (column 1) and meal logging (column 3) on treatment assignment. Columns 2 and 4 show the differences between expected and actual rates of meditation and meal logging, respectively. Participants in the control group grossly over-predict their rates of both behaviors: they expect to meditate 38% of the time when actually they do so 9.4% of the time, and they expect to log their meals 52% of the time when they actually do so 11.8% of the time. Participants who receive only meditation messages predict their meditation rates to be even higher than the control group, but the level of over-prediction is similar, resulting in an "expected target effect" (11%) that is actually quite close to the true target effect (8.8%). Participants who receive only nutrition messages also predict their meal-logging rates to be higher than the control group, but here they over-predict less, resulting in an "expected target effect" (12%) that is significantly lower than the true target effect (16.6%).

"Expected spillover effects" of nutrition messages and incentives on meditation are small and positive (though neither of these effects are significant). Participants thus significantly over-estimate spillover effects by 5 percentage points in both cases (i.e. they under-estimate negative spillovers by 5 percentage points). In the case of the spillover of meditation messages on meal logging, we cannot conclude that participants correctly predicted the negative spillover (the expected effect is about 70% of the true spillover, at -0.017, but the standard error is high), but there is no evidence that they under-estimated spillovers either. These results are consistent with recent evidence of low sophistication regarding one's limited attention (Bronchetti et al., 2020).

4.4 Opt-Out Behavior

To opt-out, participants simply had to reply "STOP" to the same number from which they were receiving messages. They were informed of this in the consent form, by email upon treatment assignment, as well as in the first text message they received. Table 7 examines whether participants receiving meditation messages ever opted out (column 1), and whether participants receiving nutrition messages ever opted out (column 2), where the omitted group has one set of messages, and the treatment group has both. Because the analysis in each column is restricted to those who received a given set of messages (m_x in column 1, m_y in column 2), I estimate only one effect: that of having the additional set of messages. Because I do not have measures of opt-out for control participants, I cannot estimate the interaction term, or "interference" (I can only estimate the equivalent of the linear combinations reported in Table 3, $m_x + m_x m_y$ and $m_y + m_x m_y$).

Baseline levels of opt-out were relatively high, at 14.5% for the meditation-only group and 16.8% for the nutrition-only group. There is no evidence that participants receiving both sets of messages opted out more frequently than those with just one. For opt-out of the meditation program, the coefficient has the opposite sign. For opt-out of the nutrition program, the coefficient has the correct sign but is not statistically different from zero ($p = 0.11$).

Thus, although receiving an additional set of non-target messages did reduce rates of the target behavior relative to the group that received only target messages ($m_x + m_x m_y < 0$ and $m_y + m_x m_y < 0$ in Table 3), it did not induce more opt-out behavior. If it did, it would hint at the possibility of overload, since the non-target messages clearly interfere with the target messages in causing people to opt out of them.

4.5 Reading and Recall of Messages

To understand the extent to which participants actually read and recalled the content of messages, I use two additional measures that I collected during the experiment. The first measure is whether or not participants responded to a surprise raffle sent to participants with messages via SMS.[38] The message said: "Hi from Remindful/eNOMerate! We are offering a surprise raffle for a USD 20 Amazon gift card. To enter, tap [link] and press send. Msg & dta rates may apply." Each message participant received a maximum of one raffle message over the course of the experiment. Roughly half of message participants received the raffle on day 10 or 11 (halfway through the second week) and the other half received the raffle on

[38] Due to an implementation error, we are missing this data from 592 participants with messages. These participants accidentally received messages with a broken link, and so we do not know whether they responded or not.

day 20 or 21 (at the end of the third week). Participants receiving both messages were randomly assigned to receive either the eNOMerate raffle or the Remindful raffle.

Table 8, displays the results in columns 1 and 3. Again, because the analysis in each column is restricted to those who received a given set of messages (m_x in column 1, m_y in column 3), I estimate only the effect of having the additional set of messages. I find that raffle response rates were low even in the groups with one set of messages, at 31% and 26% for meditation and meal logging raffles, respectively (but higher conditional on not opting-out, at 36% and 31% respectively). The group with both sets of messages was about 30% less likely to respond to both raffles, suggesting that they were reading messages at a lower rate.

The second indicator is knowledge about meditation and nutrition, as measured by the percentage of questions answered correctly on a quiz administered at the end of the study, one month after the end of treatment. The quiz consisted of true/false questions on information provided in the messages, with additional options to answer "I remember seeing this message but I do not remember the details" or "I do not remember seeing this message." Participants received 1 point for every correct answer, 0 point for every incorrect or "I do not remember seeing this message" answer, and 0.5 points for answering "I remember seeing this message but I do not remember the details." They were unaware of this scoring system, and had no explicit incentives to perform well. Both raffle response rates and quiz scores are restricted to participants with messages (I do not quiz groups on information they did not receive), so the omitted group has one set of messages, and the treatment group will have both.

The group with both sets of messages may have done slightly worse on the knowledge quizzes, as demonstrated in columns 2 and 4 of Table 8, but these effects are small and not significant. The coefficient of interest does not change substantially when I restrict the sample to those who did not opt out, suggesting that the bulk of the effect is driven by participants who continued to receive messages throughout the treatment period.

The fact that participants with both sets of messages are less likely to respond to a surprise message when they have an incentive to do so supports the possibility that non-target messages induced a negative spillover on the reading of target messages. Recall from Table 3 that participants with both sets of messages are 2.2 percentage points less likely to meditate than those with meditation messages alone, and 5.0 percentage points less likely to log their meals than those with meal-logging messages alone. Thus a negative spillover on reading would be consistent with, and potentially contributes to, the negative spillovers on behavior that we observe.

Interestingly, the negative spillover on reading is also indicative of overload. Again, because I do not have these measures for control participants, I cannot estimate interference or the interaction between the two sets of messages. But recall that overload was introduced to capture the possibility that non-target messages interfere with target messages, making them less effective (see Figure A1). If non-target messages induced negative spillovers on the reading of target messages, this is tantamount to interference with those messages, or overload. These data thus offer an alternative measure of overload that was not captured by the simple framework.

How is this consistent with the fact that we did not find evidence of interference with respect to behavior? There are several possibilities. First, as discussed above, interference may not be a good test for overload in this context, given that the requisite assumption of symmetric overload may not hold. Second, if message reading only weakly translates into behavior, then we had more statistical power to detect interaction effects with respect to message reading.[39] Third, suppose the messages worked mostly as prompts, and not as information. Then just seeing or hearing the notification (and not actually reading the message) may have been sufficient to induce action and create the treatment effects we observe. Then, if prompts do not generate overload, but information does, we would expect to see overload with respect to the reading of messages, but not with respect to behavior, which is exactly what we see.

[39] We might also worry that the failure to find evidence of interference masks important heterogeneity across subjects. For example, we might expect interference effects to differ across participants with different amounts of *total* notifications. (The effect could go in either direction, and in Appendix I I provide an example of how my model might be modified to generate predictions about heterogeneous spillover effects by total notifications.) In Appendix I I estimate treatment effects separately for participants with baseline notifications above and below the median. I find no evidence of heterogeneity by the baseline amount of notifications, but it should be noted that the experiment was not powered to detect these effects.

5 Structural Estimation

In the section that follows, I parameterize and estimate the model from Section 2. The purpose of this estimation is to quantify the contribution of each mechanism to the observed spillovers (Section 5.3).

5.1 Structural Model

Let $C(a_x, a_y, w_x, w_y) = f(a_x, a_y) - s^x(w_x, w_y)a_x - s^y(w_x, w_y)a_y$, where $f(a_x, a_y) = \frac{1}{2}\alpha(a_x^2+a_y^2)+\rho a_x a_y$. Let the effort subsidy in the case of x messages be $s^x(m_x, m_y) = \nu_x m_x + \gamma m_x m_y + \theta_m m_y$, and in the case of y messages, $s^y(m_y, m_x) = \nu_y m_y + \gamma m_x m_y + \theta_m m_x$. Let the effort subsidy in the case of incentives be $s^x(z_x, z_y) = \lambda z_x + \theta_z z_y$ for x and $s^y(z_y, z_x) = \lambda z_y + \theta_z z_x$ for y. This means that $c^x = \alpha a_x + \rho a_y - (\nu_x m_x + \gamma m_x m_y + \theta_m m_y)$ and $c^y = \alpha a_y + \rho a_x - (\nu_y m_y + \gamma m_x m_y + \theta_m m_x)$.

I thus assume that subsidies targeting behavior j reduce the marginal cost of effort to j by a fixed amount s^j. I allow target messages to have different effort subsidies depending on the behavior (allowing different ν_x and ν_y) but I assume that non-target messages impose the same tax regardless of the behavior (fixed θ_m, γ for x and y). I also impose that $c_1^x = c_1^y$.[40] In this parameterization, ρ captures depletion (net of complementarity), θ_m and θ_z capture diversion for messages and incentives, respectively, and γ captures overload. In my experiment I will only have incentives for y, so z_x will always be zero and λ will represent the target effect of incentives for y.

In each period, I can thus write the agent's problem as:

$$\max_{a_x, a_y} \Big\{ a_x u_x + a_y u_y - \Big(\frac{1}{2}\alpha(a_x^2 + a_y^2) + \rho a_x a_y \\ - a_x(\nu_x m_x + \gamma m_x m_y + \theta_m m_y + \theta_z z_y) - a_y(\nu_y m_y + \gamma m_y m_x + \lambda z_y + \theta_m m_x)\Big)\Big\}$$

Let a_{xit}^* represent the optimal attention paid to behavior x by individual i in period t, with a_{yit}^* defined similarly. I allow for individual heterogeneity in u_x and u_y. I let $u_{xi} = \mu_x + \epsilon_{xi}$, and I normalize μ_y to 1, so that $u_{yi} = 1 + \epsilon_{yi}$. I assume that ϵ_{xi} and ϵ_{yi} are jointly normal, with mean zero, variances $\sigma^2_{\epsilon_x}$ and $\sigma^2_{\epsilon_y}$, and covariance $\sigma_{\epsilon_x \epsilon_y}$. I also allow for individual heterogeneity in the effects of

[40] I allow x and y to have different baseline returns (μ_x) and responses to target messages (ν_x, ν_y). All other differences between x and y (including any differences between c_1^x and c_1^y) are not identified, and will be loaded onto the aforementioned parameters.

target messages, so that $\nu_x = \phi_x + \delta_{xi}$ and $\nu_y = \phi_y + \delta_{yi}$. I assume that δ_{xi} and δ_{yi} are jointly normal, with mean zero, variances $\sigma^2_{\delta_x}$ and $\sigma^2_{\delta_y}$, and covariance $\sigma_{\delta_x \delta_y}$. I define $a^*_{xi} = E_{\xi_x}[a^*_{xit}]$, which I can estimate in the data as: $\hat{a}^*_{xi} = \frac{1}{T}\sum_{t=1}^{T} x_{it}$. I define a^*_{yi} and $\hat{a}^*_{yi}$ similarly. This means that my estimation is based on 3,845 participant-level observations of average meditation and meal logging rates over the 27 days of the treatment period.

In the benchmark specification I will allow θ (diversion) to differ across messages m and incentives z, but not across behaviors x and y. I also report estimates of the alternative specification, in which I allow θ to differ across behaviors but not across intervention type, in Appendix Table A5.

5.2 Estimation

I estimate the 15 parameters of the model using classical minimum distance. I use all control means and treatment effects on a_x, a_y, $var(a_x)$, $var(a_y)$, and $cov(a_x, a_y)$, which amounts to 22 moments, written out in Appendix E.[41] Let ζ represent the vector of q parameters and let $m(\zeta)$ represent the r moments as functions of the parameters. The minimum-distance estimator selects parameters $\hat{\zeta}$ that minimize the expression $(m(\zeta) - \hat{m})'W(m(\zeta) - \hat{m})$.[42,43]

The simple moment conditions make it relatively straightforward to see how parameters are identified. Most important is the identification of the relationship between α and ρ, which will allow us to distinguish depletion from diversion. This relationship is identified by the four target and spillover effects. The intuition is the same as that described in Section 4.1: if ρ is zero, then we expect the spillovers generated to be the same across all interventions, since θ_m has been assumed not to differ between x and y. If ρ is positive but θ_m is zero, then the difference between spillovers generated by m_x versus those generated by m_y should reflect the difference between their target effects, since both are driven by differences between ϕ_x

[41] I do not use the effects of incentives on variances and covariances because for the sake of simplicity, I have not allowed for heterogeneity in the incentive attention subsidy.

[42] For the weighing matrix W I use the diagonal of the inverse of the variance-covariance matrix of the moments. With small samples, using the full variance-covariance (VC) matrix results in biased estimates (Altonji and Segal, 1996). Indeed, in this case, similar estimates are obtained using the diagonal of the VC matrix or the identity matrix, while different estimates are obtained using the full VC matrix. I show estimates using the identity matrix in Appendix F.

[43] I estimate the variance of $\hat{\zeta}$ as $(\hat{G}'W\hat{G})^{-1}(\hat{G}'W\hat{\Lambda}W\hat{G})(\hat{G}'W\hat{G})^{-1}$, where $\hat{G} = \Delta_\zeta m_n(\hat{\zeta})$ (the matrix of derivatives of the moments with respect to parameters, evaluated at the estimated parameters) and $\hat{\Lambda} = Var(\hat{m})$. Importantly, these standard errors do not take into account model uncertainty.

and ϕ_y. The relationship between α and ρ is also identified by the differences in the covariance between a_{xi} and a_{yi} across treatments. Again, the intuition is the same as that described in Section 4.1: if ρ is zero, then there should be no difference in the covariance across treatments; if ρ is positive, then treatments that induce higher x or y should also induce a lower covariance.

Once I have pinned down the relationship between α and ρ, I can use the control group means to separately identify α and ρ. The remaining parameters are straightforward to identify once α and ρ are known. Specifically, ϕ_x, ϕ_y, λ, θ_m, and θ_z are identified by the three target effects and three spillover effects of messages and incentives. The identification of the diversion parameters depends critically on ρ having already been pinned down. γ is (over-)identified by the two interference effects; this arises mechanically from the way overload was defined in the model. σ_{ϵ_x}, σ_{ϵ_y}, and σ_{xy} are identified by the variances and covariances of a_{xi} and a_{yi} in the control group, and σ_{phi_x}, σ_{phi_y}, and $\sigma_{\phi_x\phi_y}$ are identified by the variances and covariances of a_{xi} and a_{yi} across treatments.

The maximized value of the objective function is asymptotically distributed as $\chi^2(r-q)$, so the critical value for an over-identification test of model fit is 2.17. The maximized value of the objective function is 12.3, so the test is rejected. Figure A4 compares the actual moments and predicted moments. The high test statistic is driven principally by moment 12, the effect of m_y on the variance of a_x. It turns out that in the data, this estimate is significantly negative, at -0.015 (0.005). In the model, the variance is indeed predicted to fall, but only by a very small amount.

I report parameter estimates in Table 9. There is not strong evidence of overload, as $\hat{\gamma}$ is negative but not significantly different from zero. We cannot say much about depletion (net of complementarity), as $\hat{\rho}$ is positive but its standard error is very high. Message diversion is negative and statistically significant; incentive diversion is closer to zero and has a much higher standard error. I conduct various robustness checks in Appendix F.[44],[45],[46]

[44]In Appendix Table A4, I present estimates using the same moments and parameters, but using the identity matrix as the weighing matrix. The estimates are very similar and the main conclusions are consistent.

[45]In Appendix Table A5, I present estimates using the same moments but replacing parameters θ_m and θ_z with θ_x and θ_y. The main conclusions are consistent.

[46]In Appendix Figure A5 I examine the sensitivity of the four key parameter estimates to moments as in Andrews et al. (2017).

5.3 Decomposing Spillovers

I can now use my parameter estimates to decompose the observed spillovers. Note the expressions for the spillover moments that come out of the structural model (where $\omega = \alpha^2 - \rho^2$):

$$\mathbb{E}[a_x^*|m_x = 0, m_y = 1] - \mathbb{E}[a_x^*|m_x = 0, m_y = 0] = \frac{\alpha\theta_m - \rho\phi_y}{\omega} \tag{5}$$

$$\mathbb{E}[a_y^*|m_x = 1, m_y = 0] - \mathbb{E}[a_y^*|m_x = 0, m_y = 0] = \frac{\alpha\theta_m - \rho\phi_x}{\omega} \tag{6}$$

In each expression, the first term represents the component of the spillover driven by diversion, and the second term represents the component of the spillover driven by depletion. Using my parameter estimates, and bootstrapping standard errors, I find that diversion explains 85.6% of spillovers generated by m_y on x (S.E. 67), and 92.3% of spillovers generated by m_x on y (S.E. 32). The standard errors are large, but for the latter, a 95% confidence interval does not contain zero. These estimates can be understood as a quantification of the reduced form evidence, presented in Section 4.1, that diversion is a necessary ingredient of the spillovers I observe.

What can we say about the presence or absence of depletion? Recall that ρ potentially encompasses three things: time constraints, costly cognitive effort, and complementarity between behaviors x and y. There are thus several forces that might push ρ upwards. The time constraint might be binding, despite the choice of non-time-consuming behaviors, or people might create artificial time constraints within mental accounts that do bind (i.e. allotting 30 minutes per day for wellness). Perhaps moral licensing causes negative complementarity between these two positive health behaviors in the utility function. The fact that the experiment artificially bundled these two behaviors might cause negative complementarity between the behaviors (I discuss this in detail in Section 7 below). Or, finally, engaging in these behaviors might require costly mental effort that depletes internal resources. There are also several forces that might push ρ downwards. There could be positive complementarity in the utility function between behaviors x and y: meditation might cause people to care more about their health, raising the utility from logging meals. It is also possible that the fact that the experiment artificially bundled these two behaviors created a reminder effect, if the act of meditating also reminded people to log their meals, reducing the mental effort of doing so. These possibilities, as well as the high standard error around the estimate of ρ, make it difficult to say anything about depletion.

6 Implications for Cost Effectiveness Analysis

My results have two implications for cost-effectiveness analysis. First, the existence of spillovers implies that interventions are costlier than they appear. Second, recall that spillovers driven by diversion, unlike those driven by depletion, can exist even when the intervention does not induce the target behavior. This means, as we saw with the meal logging incentives and messages, that high-impact interventions will not necessarily create larger spillovers than low-impact ones. More generally, there may be unappreciated differences in cost-effectiveness across interventions.

In this section, I will conduct a simple cost-effectiveness analyses to quantify these implications, exploiting the comparison between interventions m_y and z_y.[47] Suppose a policymaker cares about the total cost-effectiveness of a set of interventions, taking multiple behaviors into account. Let the total effect of a set of interventions on multiple behaviors be simply the sum of the normalized effects on the set of behaviors. In our context, the total effect of a set of interventions will be equal to their net effect on meditation (in standard deviations), plus their net effect on meal logging (in standard deviations). I will consider two sets of interventions. Policy A consists of messages about both x and y, and costs $2p_m$, where p_m is the cost of one message set. Policy B consists of messages about x and incentives for y, and costs $p_m + p_z$, where p_z is the cost of incentives. I will assume that there are no interaction effects between interventions.[48]

I will use the following facts from the reduced form analysis. For meditation, the standard deviation in the control group was 0.29. This means that m_x, generated a target effect of 0.30 S.D.'s (0.088) and a spillover effect of -0.08 S.D.'s (-0.024). For meal logging, the standard deviation in the control group was 0.32. This means that m_y generated a target effect of 0.52 S.D.'s (0.166) and a spillover effect of -0.09 S.D.'s (-0.028); and z_y generated a target effect of 1.19 S.D.'s (0.381) and a spillover

[47] Ideally, I would make use of the model to simulate counterfactual policies, and analyze their cost-effectiveness. In practice, since I did not randomly vary the intensity of interventions, this would require too many assumptions.

[48] This is indeed the case for m_x and m_y. I did not measure it for m_x and z_y but I believe it is a reasonable assumption, given that the hypothesized mechanism for interference was information overload, which z_y is unlikely to induce.

effect of -0.08 S.D.'s (-0.025).

In a hypothetical world without spillovers, Policy A would achieve a total effect of $0.3 + 0.52 = 0.82$ S.D.'s, and Policy B would achieve a total effect of $0.3 + 1.19 = 1.39$ S.D.'s. But with spillovers, Policy A achieves a total effect of $0.3 + 0.52 - 0.08 - 0.09 = 0.65$ S.D.'s, and Policy B achieves a total effect of $0.3+1.19-0.08-0.08 = 1.33$ S.D.'s. This means that under Policy A, which costs $2p_m$, the cost per standard deviation rises from $2p_m/0.82 = 1.22p_m$ to $2p_m/0.65 = 1.54p_m$ once spillovers are taken into account, or by 26%. In contrast, under Policy B, which costs p_z, the cost per standard deviation rises from $(p_m+p_z)/1.39 = 0.72p_z$ to $(p_m+p_z)/1.33 = 0.75p_z$ once spillovers are taken into account, or by 4.2%. The costs of spillovers are thus substantial for Policy A, but negligible for Policy B, which generates a large effect on y without a correspondingly large spillover effect on x.

In order to compare the cost-effectiveness of Policy A and Policy B, let's ask what p_m and p_z would need to be in order for the total cost-effectiveness of Policy A to equal that of Policy B. In the absence of spillovers, Policy A has cost per standard deviation of $2p_m/0.82$, and Policy B has cost per standard deviation of $(p_m + p_z)/1.39$, so they are equally cost-effective when $p_z = 2.4p_m$. This means that Policy B will be more cost-effective as long as $p_z < 2.4p_m$. In the presence of spillovers, Policy A has cost per standard deviation of $2p_m/0.65$, and Policy B has cost per standard deviation of $(p_m + p_z)/1.33$, so they are equally effective when $p_z = 3.1p_m$. This means that Policy B will be more cost-effective as long as $p_z < 3.1p_m$. In the experiment conducted here, incentives cost $5.23 per capita, and messages cost $1.89 per capita, so $p_z = 2.8p_m$. Failing to consider spillovers would thus have resulted in the choice of the less cost-effective policy. More generally, in the presence of spillovers, we see a shift in the relative cost-effectiveness of different interventions. Specifically, high-impact interventions might be more cost-effective than low-impact ones once multiple behaviors are considered.

7 Alternative Explanations

7.1 Can time constraints account for spillovers?

As discussed in Section 5, the parameter ρ encompasses several things: time constraints, costly cognitive effort, and complementarity between x and y. The fact that the estimate for ρ was not significantly different from zero thus tells us little about the importance of time constraints. What can we conclude, then? First, most basically, the results show that even behaviors that take very little time can generate spillovers on other behaviors. Second, the results suggest that ρ (and therefore time constraints) cannot *alone* explain the observed spillovers. As an additional check of the second conclusion, in Appendix G I use my estimates to compute the implied elasticity of substitution between meditation and meal logging in the case that spillovers are driven fully by the time constraint.[49] I find that in order to explain the observed spillovers, the elasticity of substitution would need to be at least 1331. In summary, these behaviors take so little time that in a standard model, the elasticity of substitution would need to be unrealistically high in order for the time constraint to explain the observed spillovers.

7.2 Can the bundling of treatments or experimenter demand effects account for spillovers?

In this section I address two alternative explanations for spillovers. The first concerns the fact that interventions in the real world are not typically bundled the way they were in this experiment, and that the bundling of the two behaviors might have encouraged a trade-off mindset or mental account that led to spillovers. The experiment design attempts to address this concern in two ways. First, in the baseline survey, I present the experiment as a study on "wellness behaviors," and every time I ask about meditation or meal logging, I also ask about sleep and exercise. Moreover, at the end of the baseline survey when the assignment of treatments is explained, participants are told that, "we'll randomly assign you to one or several (or none) of our messaging or incentive programs for one or several wellness behaviors," so participants have no reason to believe that treatments are limited to meditation and meal logging (as they ultimately are). Thus, the only sense in which meditation and meal logging were "bundled" together separately from exercise and sleep was through the apps (participants were only asked to download two apps, about meditation and meal logging, not four). Second, as discussed in Section 3, I give the meditation and meal logging programs separate names and send messages from separate numbers.

The second, related concern is that spillovers might be driven by a "negative" experimenter demand effect. If participants were not explicitly encouraged to do

[49]This exercise does not account for more sophisticated models of time constraints within mental accounts (i.e. allotting 30 minutes a day to wellness).

a behavior, they might feel as if they were “not supposed” to do it. For example, someone who was randomly assigned to receive messages about meditation might have believed that they were not supposed to log their meals. This was indeed a concern: in piloting, several participants expressed confusion about what they were “supposed” to do. For this reason, when we rolled out the experiment (and as described in Section 3), we added to the enrollment confirmation email, ”Regardless of your treatment assignment, your ultimate use of the apps is entirely up to you” in attempt to make it clear that participants were not “supposed” to do anything in particular for the purpose of the experiment or experimenter.

Did these measures succeed in ensuring the behaviors were not perceived as bundled, and in inhibiting any experimenter demand effects? In Survey 2, in addition to asking the extent to which people “expected” to do each behavior, I also asked the extent to which they “hoped” to do each behavior. Moreover, I asked about exercise and sleep, in addition to meditation and meal logging. If participants bundled meditation and meal logging in a way that resulted in a “trade-off mindset” and drove spillovers, then being assigned a treatment for one behavior should lead them to “hope” less that they would do the other. Similarly, if participants felt that being assigned a treatment for one behavior meant that they “were not supposed” to do the other, then we would also expect them to have less “hope” of doing it. In Panel B of Appendix Table A6 I show the effects of treatment assignment on self-reported hopes (how many days per week participants “hope” to do the behavior, divided by 7). I find no evidence that participants assigned to meditation or meal logging treatments are less hopeful of doing the alternative behavior. Indeed, participants assigned to meal logging messages actually “hoped” to meditate more than the control group. I can also compare the spillover effects of m_x, m_y and z_y on hopes regarding the non-target “bundled” behavior (meditation, meal-logging), to hopes regarding the non-target ”non-bundled” behaviors (exercise, sleep). This amounts to 6 pairwise comparisons. I find just one significant negative difference: those assigned to m_x are less hopeful to meal log than they are to exercise. In sum, there is little evidence that participants “hoped” to meditate or meal log less because they were assigned to the other bundled behavior.

7.3 Are the results specific to this context?

Above and beyond the bundling of treatments, we might worry that the observed spillovers are specific to low-stakes behaviors, to health and wellness more generally, or to the population of wellness-loving Facebook users. This is, of course, possible. However, as briefly discussed in the introduction, the context, while specific, was strategically chosen on two dimensions. First, it is relevant: nudges for small-scale health-related behavior change are increasingly used, at scale, by corporations and governments across the globe. 69% of U.S. adults use Facebook (Pew Research Center, 2021), and the individuals that selected into my experiment are arguably the ones who would be compliers in more traditional experiments. This may not be ideal for classic program evaluation, where the intent-to-treat estimate might be more policy-relevant, but it is beneficial in a study focused on spillovers. Second, this is a context where neither standard economic models nor beneficiaries themselves (see Table 6) anticipate spillovers. Given the sparsity of empirical research on this question, studying a context with widespread nudges but without strong priors regarding spillovers is a valuable starting point, even if the stakes, in this case, are low.

8 Conclusion

A growing literature has documented a wide variety of interventions that shift behavior and generate meaningful economic impacts. This paper has demonstrated the importance of studying these interventions in a broader context, taking into account how they affect other interventions and other behaviors. Namely, costly mental effort and attention can cause interventions to have negative spillovers on other behaviors in unexpected contexts. Spillovers are driven in large part by diversion, which implies that they will not necessarily grow with the effectiveness of the intervention. This paper raises many questions about the generalizability of these effects to other behaviors, contexts, and interventions.

References

Afshin, A., Sur, P. J., Fay, K. A., Cornaby, L., Ferrara, G., Salama, J. S., Mullany, E. C., Abate, K. H., Abbafati, C., Abebe, Z., et al. (2019). Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*.

Agarwal, S., Driscoll, J. C., Gabaix, X., and Laibson, D. (2013). Learning in the credit card market. Working paper, National Bureau of Economic Research.

Allcott, H. and Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, 11(1):236–76.

Allcott, H., Lockwood, B. B., and Taubinsky, D. (2019). Regressive sin taxes, with an application to the optimal soda tax. *The Quarterly Journal of Economics*, 134(3):1557–1626.

Allcott, H. and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10):3003–37.

Altmann, S., Grunewald, A., and Radbruch, J. (2019). Passive choices and cognitive spillovers. *Review of Economic Studies*, Forthcoming.

Altonji, J. G. and Segal, L. M. (1996). Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics*, 14(3):353–366.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics*, 132(4):1553–1592.

Bénabou, R. and Tirole, J. (2004). Willpower and personal rules. *Journal of Political Economy*, 112(4):848–886.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Skimmyhorn, W. L. (2021). Borrowing to save? The impact of automatic enrollment on debt. Working paper, SSRN.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, 127(3):1243–1285.

Botvinick, M. and Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annual Review of Psychology*, 66:83–113.

Brandon, A., List, J. A., Metcalfe, R. D., Price, M. K., and Rundhammer, F. (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences*, 116(12):5293–5298.

Bronchetti, E. T., Kessler, J. B., Magenheim, E. B., Taubinsky, D., and Zwick, E. (2020). Is Attention Produced Rationally? Working paper, National Bureau of Economic Research.

Burgess, P. W., Scott, S. K., and Frith, C. D. (2003). The role of the rostral frontal cortex (area 10) in prospective memory: A lateral versus medial dissociation. *Neuropsychologia*, 41(8):906–918.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2019). Measuring the welfare effects of shame and pride. *American Economic Review*, Forthcoming.

Chernev, A., Böckenholt, U., and Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25(2):333–358.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., and Olsen, T. (2014). Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark. *The Quarterly Journal of Economics*, 129(3):1141–1219.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77.

Chun, M. M., Golomb, J. D., and Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual review of psychology*, 62:73–101.

Damgaard, M. T. and Gravert, C. (2018). The hidden costs of nudging: Experimental evidence from reminders in fundraising. *Journal of Public Economics*, 157:15–26.

DellaVigna, S., List, J. A., and Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics*, 127(1):1–56.

Deryugina, T. and Marx, B. M. (2021). Is the Supply of Charitable Donations Fixed? Evidence from Deadly Tornadoes. *American Economic Review: Insights*, 3(3):383–98.

Dolan, P. and Galizzi, M. M. (2014). Because I'm worth it: a lab-field experiment on the spillover effects of incentives in health.

Dolan, P. and Galizzi, M. M. (2015). Like ripples on a pond: behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, 47:1–16.

Donkers, B., van Diepen, M., and Franses, P. H. (2017). Do charities get more when they ask more often? Evidence from a unique field experiment. *Journal of behavioral and experimental economics*, 66:58–65.

Dubois, P., Griffith, R., and O'Connell, M. (2020). How well targeted are soda taxes? Technical Report 11.

Duncan, J., Martens, S., and Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, 387(6635):808.

Dunn, T. L., Lutes, D. J., and Risko, E. F. (2016). Metacognitive evaluation in the avoidance of demand. *Journal of Experimental Psychology: Human Perception and Performance*, 42(9):1372.

Edmunds, A. and Morris, A. (2000). The problem of information overload in business organisations: A review of the literature. *International Journal of Information Management*, 20(1):17–28.

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., and Walker, M. W. (2019). General equilibrium effects of cash transfers: Experimental evidence from Kenya. Working paper, National Bureau of Economic Research.

Ek, C. and Miliute-Plepiene, J. (2018). Behavioral spillovers from food-waste collection in Swedish municipalities. *Journal of Environmental Economics and Management*, 89:168–186.

Embi, P. J. and Leonard, A. C. (2012). Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. *Journal of the American Medical Informatics Association*, 19(e1):e145–e148.

Eppler, M. J. and Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society*, 20(5):325–344.

Farhi, E. and Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review*, 110(1):298–336.

Fawcett, J., Kingstone, A., and Risko, E. (2015). *The Handbook of Attention*. MIT Press.

Fricke, H., Kalogrides, D., and Loeb, S. (2018). It's too annoying: Who drops out of educational text messaging programs and why. *Economics letters*, 173:39–43.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710.

Galizzi, M. M. and Whitmarsh, L. (2019). How to measure behavioral spillovers: A methodological review and checklist. *Frontiers in Psychology*, 10:342.

Gonen-Yaacovi, G. and Burgess, P. (2012). Prospective memory: The future for future intentions. *Psychologica Belgica*, 52(2-3).

Goyal, M., Singh, S., Sibinga, E. M., Gould, N. F., Rowland-Seymour, A., Sharma, R., Berger, Z., Sleicher, D., Maron, D. D., Shihab, H. M., et al. (2014). Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. *JAMA Internal Medicine*, 174(3):357–368.

Hagmann, D., Ho, E. H., and Loewenstein, G. (2019). Nudging out support for a carbon tax. *Nature Climate Change*, 9(6):484–489.

Hall, J. D. and Madsen, J. (2020). Can behavioral interventions be too salient? Evidence from traffic safety messages. Working paper, SSRN.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–82.

Hanna, R., Duflo, E., and Greenstone, M. (2016). Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves. *American Economic Journal: Economic Policy*, 8(1):80–114.

Hanna, R., Mullainathan, S., and Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3):1311–1353.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(4):1009–1055.

Hussam, R. and Oh, D. (2021). Behavioral Transmission: Evidence from a Public Health Campaign in Bangladesh. Working paper.

Hwang, M. I. and Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science*, 25(3):213–218.

Institute for Health Metrics and Evaluation (2018). Findings from the Global Burden of Disease Study 2017. Technical report, Seattle, Washington.

Jessoe, K., Lade, G., Loge, F., and Spang, E. (2017). Spillovers from behavioral interventions: Experimental evidence from water and energy use.

Jimenez-Gomez, D. (2018). Nudging and phishing: A theory of behavioral welfare economics. *Available at SSRN 3248503*.

Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, 62(12):3393–3411.

Khan, U. and Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research*, 43(2):259–266.

Lacetera, N., Pope, D. G., and Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, 102(5):2206–36.

Luck, S. J. and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657):279.

Medina, P. C. (2020). Side Effects of Nudging: Evidence from a Randomized Intervention in the Credit Card Market. *The Review of Financial Studies*, 34(5):2580–2607.

Meer, J. (2017). Does fundraising create new giving? *Journal of Public Economics*, 145:82–93.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. *Proceedings of the National Academy of Sciences*, 108(26):10415–10420.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2):81.

Morrison, W. and Taubinsky, D. (2019). Rules of thumb and attention elasticities: Evidence from under- and over-reaction to taxes. *Review of Economics and Statistics*, Forthcoming.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. (2017). General equilibrium effects of (improving) public employment programs: Experimental evidence from India. Working paper, National Bureau of Economic Research.

Nafziger, J. (2020). Spillover effects of nudges. *Economics Letters*, 190:109086.

Neisser, U. and Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive psychology*, 7(4):480–494.

O'Donoghue, T. and Rabin, M. (2006). Optimal sin taxes. *Journal of Public Economics*, 90(10-11):1825–1849.

Persson, P. (2018). Attention manipulation and information overload. *Behavioural Public Policy*, 2(1):78–106.

Pew Research Center (2021). Social media use in 2021. Technical report, Washington, D.C.

Rogers, T. and Milkman, K. L. (2016). Reminders Through Association. *Psychological Science*, 27(7):973–986.

Scharf, K. A., Smith, S., and Wilhelm, M. (2017). Lift and shift: the effect of fundraising interventions in charity space and time. Technical report.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., and Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, 40:99–124.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Spiegler, R. (2015). On the equilibrium effects of nudging. *The Journal of Legal Studies*, 44(2):389–416.

Taubinsky, D. (2013). From intentions to actions: A model and experimental evidence of inattentive choice. Working paper.

Taubinsky, D. and Rees-Jones, A. (2018). Attention variation and welfare: theory and evidence from a tax salience experiment. *The Review of Economic Studies*, 85(4):2462–2496.

Taylor, R. L. (2020). A mixed bag: The hidden time costs of regulating consumer behavior. *Journal of the Association of Environmental and Resource Economists*, 7(2):345–378.

Thaler, R. H. and Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Tiefenbeck, V., Staake, T., Roth, K., and Sachs, O. (2013). For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy*, 57:160–171.

Wertenbroch, K. (1998). Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Science*, 17(4):317–337.

Wharton, C. M., Johnston, C. S., Cunningham, B. K., and Sterner, D. (2014). Dietary self-monitoring, but not dietary quality, improves with use of smartphone app technology in an 8-week weight loss trial. *Journal of Nutrition Education and Behavior*, 46(5):440–444.

Yantis, S. and Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *Journal of Experimental Psychology: Human perception and performance*, 10(5):601.

Zhou, F., Shefer, A., Wenger, J., Messonnier, M., Wang, L. Y., Lopez, A., Moore, M., Murphy, T. V., Cortese, M., and Rodewald, L. (2014). Economic evaluation of the routine childhood immunization program in the United States, 2009. *Pediatrics*, 133(4):577–585.

Table 1: Experiment Design

Group	Description	Messages m_x = 1	m_y = 1	Incentives z_y = 1	N
1	ctrl				814
2	x messages	x			763
3	y messages		x		803
4	x & y messages	x	x		822
5	y incentives			x	643
					3845

Notes: The experiment included five treatment groups: a control group, a group that received only messages about meditation (x), a group that received only messages about meal logging (y), a group that received messages about both meditation (x) and meal logging (y), and a group that received incentives to log meals (y).

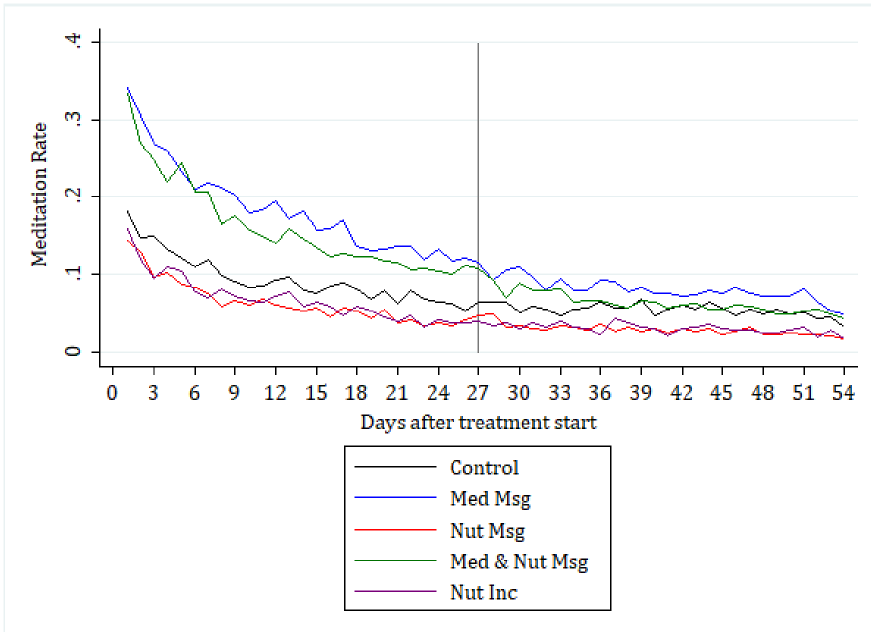
Table 2: Orthogonality Check

	control	mx	my	mx & my	zy	F-test, joint sig
female†	0.93	0.93	0.93	0.93	0.93	1.00
	0.25	0.25	0.25	0.25	0.26	
went to college†	0.71	0.72	0.71	0.71	0.72	0.99
	0.45	0.45	0.45	0.45	0.45	
age	27.46	27.43	27.20	27.74	27.11	0.19
	5.72	6.05	5.22	5.48	4.90	
daily notifications†	52.59	53.32	52.70	54.16	53.40	0.99
	70.13	83.67	78.54	74.26	70.23	
meditated daily, ever	0.56	0.59	0.56	0.56	0.57	0.79
	0.50	0.49	0.50	0.50	0.49	
meditated daily, last month†	0.47	0.46	0.46	0.46	0.46	0.99
	0.50	0.50	0.50	0.50	0.50	
logged meals, ever	0.88	0.87	0.86	0.87	0.89	0.50
	0.33	0.34	0.35	0.34	0.32	
logged meals, last month†	0.33	0.33	0.32	0.33	0.33	0.99
	0.47	0.47	0.47	0.47	0.47	
importance, x - y	-0.44	-0.49	-0.42	-0.45	-0.47	0.99
	3.36	3.30	3.39	3.28	3.26	
difficulty - fun, x - y	-2.48	-2.76	-2.64	-2.54	-2.54	0.83
	4.93	5.07	5.35	5.16	5.10	

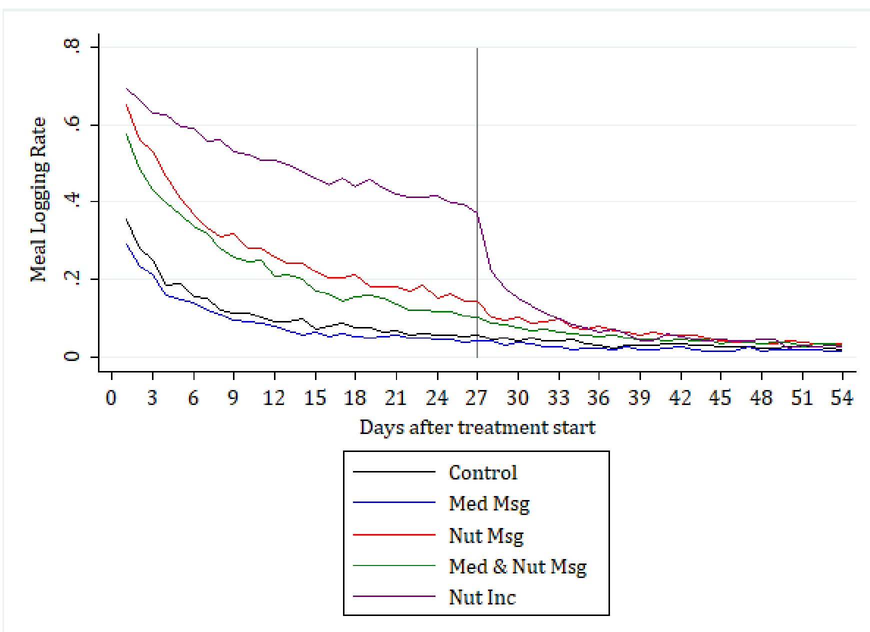
Notes: Means and standard deviations of ten variables measured in the baseline survey. † indicates variables used in the re-randomization procedure. F-test of the joint significance of treatments is reported in last column. *Daily notifications* includes all notifications the participant receives across all devices and all applications, where a notification is defined as anything that generates an alert (including SMS and email). *Importance,* $x - y$ is the "importance" of meditation, self-reported on a scale from 1 to 10, minus that of meal logging. *Difficulty,* $x - y$ is the "difficulty" of meditation, self-reported on a scale from 1 to 10, minus that of meal logging. *Fun,* $x - y$ is the "fun" of meditation, self-reported on a scale from 1 to 10, minus that of meal logging. I report only the difference between difficulty and fun, since both contribute to the experience of doing the behavior.

Figure 1: Rates of Meditation and Meal Logging by Treatment

(a) Meditation Rate by Treatment



(b) Meal Logging Rate by Treatment



Notes: Rates of meditation (a) and meal-logging (b) by treatment group, over the treatment period (days 0-27) and post-treatment period (days 28 onward). The outcome is 1 if the participant did the behavior on the given day (meditated, logged at least one meal) and 0 otherwise. This data includes the full sample of 3,845 participants. (5,845 participants were initially recruited for Survey 1; only 3,855 or 66% downloaded both apps which was a requirement for participation. Over the course of the study, 10 participants withdrew, resulting in a final sample of 3,845.)

Table 3: Reduced Form Results, Treatment Period

	Binary Outcomes		*Continuous Outcomes*	
	Meditated (0/1) (1)	Logged Meal (0/1) (2)	Min. Meditated (3)	Cal. Logged (4)
mx	0.088***	-0.024**	1.298***	-35.601**
	(0.011)	(0.010)	(0.313)	(15.110)
my	-0.028***	0.166***	-0.863***	212.193***
	(0.008)	(0.013)	(0.245)	(19.473)
mx X my	0.006	-0.026	0.432	-25.945
	(0.014)	(0.018)	(0.403)	(26.359)
zy	-0.025***	0.381***	-0.868***	350.546***
	(0.009)	(0.016)	(0.247)	(22.258)
mx + mxmy	0.094	-0.050	1.730	-61.546
	(0.009)	(0.014)	(0.252)	(21.582)
my + mxmy	-0.022	0.140	-0.430	186.248
	(0.011)	(0.012)	(0.319)	(17.736)
Ctrl Mean	0.094	0.118	2.090	153.352
Ctrl SD	0.291	0.323	8.705	473.961
Obs	102905	102905	102905	102905

Notes: OLS regressions of outcomes (binary in columns 1 and 2, continuous in columns 3 and 4) on treatments during the treatment period. mx is 1 if the individual received meditation messages and 0 otherwise; my is 1 if the individual received meal logging messages and 0 otherwise; mx X my is the interaction between mx and my, zy is 1 if the individual received meal logging incentives. Specifications include controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1\%, 5\%, and 10\% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table 4: Treatment Effects on Covariances

	Treatment Period	
	Cov(x,y) over people (within day) (1)	Cov(x,y) over time (within people) (2)
mx	0.001	0.002
	(0.004)	(0.001)
my	-0.002	0.001
	(0.004)	(0.001)
mx X my	0.021***	0.007***
	(0.005)	(0.002)
zy	-0.003	-0.001
	(0.004)	(0.001)
Ctrl Mean	0.019	0.008
Ctrl SD	0.138	0.084
Obs	102905	102905

Notes: OLS regressions at the individual-day level of the covariance between daily meditation (0/1) and daily logging of at least one meal (0/1) on treatments. Column 1 reports the covariance over people within a day (do people who meditate on a particular day also tend to log their meals?). Column 2 reports the covariance over days within a person (does someone who meditates a lot over the treatment period also log their meals a lot over the treatment period?). mx (my) is 1 if the individual received x (y) messages and 0 otherwise; mx*my is 1 if the individual received both sets of messages. The specification includes controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table 5: Reduced Form Results, Post-Treatment Period

	Binary Outcomes		*Continuous Outcomes*	
	Meditated (0/1) (1)	Logged Meal (0/1) (2)	Min. Meditated (3)	Cal. Logged (4)
mx	0.024**	-0.010	0.259	-11.675
	(0.009)	(0.006)	(0.288)	(8.646)
my	-0.025***	0.029***	-0.848***	38.311***
	(0.007)	(0.008)	(0.230)	(11.571)
mx X my	0.009	-0.002	0.613	-5.922
	(0.011)	(0.010)	(0.362)	(15.027)
zy	-0.024***	0.037***	-0.767***	35.136***
	(0.007)	(0.008)	(0.248)	(11.147)
mx + mxmy	0.034	-0.013	0.871	-17.597
	(0.007)	(0.008)	(0.220)	(12.255)
my + mxmy	-0.016	0.027	-0.235	32.389
	(0.009)	(0.007)	(0.281)	(9.576)
Ctrl Mean	0.054	0.033	1.515	41.801
Ctrl SD	0.227	0.178	7.990	253.699
Obs	102499	102499	102499	102499

Notes: OLS regressions of outcomes (binary in columns 1 and 2, continuous in columns 3 and 4) on treatments during the post-treatment period. mx is 1 if the individual received meditation messages and 0 otherwise; my is 1 if the individual received meal logging messages and 0 otherwise; mx X my is the interaction between mx and my, zy is 1 if the individual received meal logging incentives. Specifications include controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1\%, 5\%, and 10\% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table 6: Baseline Expectations of Behaviors by Treatment

	Expected Meditation (x) (1)	Expected - Actual Meditation (x) (2)	Expected Meal Logging (y) (3)	Expected - Actual Meal Logging (y) (4)
mx	0.113***	0.008	-0.017	0.009
	(0.012)	(0.015)	(0.017)	(0.018)
my	0.017	0.053***	0.122***	-0.042**
	(0.012)	(0.013)	(0.014)	(0.018)
mx X my	-0.027	-0.029	0.037	0.052*
	(0.017)	(0.021)	(0.021)	(0.025)
zy	0.021	0.053***	0.213***	-0.223***
	(0.013)	(0.014)	(0.015)	(0.021)
mx + mxmy	0.085	-0.020	0.019	0.061
	(0.012)	(0.014)	(0.013)	(0.018)
my + mxmy	-0.010	0.024	0.159	0.010
	(0.012)	(0.016)	(0.015)	(0.018)
Ctrl Mean	0.384	0.278	0.519	0.382
Ctrl Mean S.E.	(0.009)	(0.010)	(0.012)	(0.012)
Obs	2891	2876	2891	2886

Notes: OLS regressions of expected rates of behavior over treatment period, measured at baseline (Columns 1 and 3) and difference between individual's expected and actual rate (Columns 2 and 4). The specification includes controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table 7: Treatment Effects on Opting Out of Message Programs

	Opted Out Ever, Meditation Msgs (X) (1)	Opted Out Ever, Nutrition Msgs (Y) (2)
mx & my	-0.029	0.030
	(0.017)	(0.019)
mx-only Mean	0.145	
mx-only SD	0.353	
my-only Mean		0.168
my-only SD		0.374
Obs	1585	1625

Notes: OLS regressions of whether individual ever opted out of each messaging program on treatments. Restricted to individuals with a message treatment. Omitted groups are mx-only (Column 1) and my-only (Column 2). mx & my is 1 if the individual received both sets of messages. The specification includes controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month). One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table.

Table 8: Treatment Effects on Attention to Messages: Reading Rate and Memory of Content

	Raffle Response Meditation (x) (1)	Knowledge Score Meditation (x) (2)	Raffle Response Nutrition (y) (3)	Knowledge Score Nutrition (y) (4)
Panel A: Unconditional				
mx & my	-0.089**	-0.010	-0.081**	-0.008
	(0.028)	(0.006)	(0.028)	(0.007)
mx-only Mean	0.307	0.298		
mx-only SD	0.461	0.086		
my-only Mean			0.263	0.300
my-only SD			0.440	0.100
Obs	998	859	931	847
Panel B: Conditional on Not Opting-Out				
mx & my	-0.087**	-0.005	-0.084**	-0.005
	(0.033)	(0.006)	(0.034)	(0.007)
mx-only Mean	0.356	0.307		
mx-only SD	0.479	0.079		
my-only Mean			0.309	0.308
my-only SD			0.463	0.098
Obs	825	762	771	748

Notes: OLS regressions of whether individual responded to surprise raffle (Columns 1 and 3) and score on knowledge quiz (Columns 2 and 4). *Raffle response* is 1 if the individual responded to the raffle and 0 otherwise. *Score on knowledge quiz* is calculated using answers to a true/false quiz administered in the endline survey on the information provided in the messaging program. Participants received 1 point for every correct answer, 0 point for every incorrect or "I do not remember seeing this message" answer, and 0.5 points for answering "I remember seeing this message but I do not remember the details." The score is the fraction of 14 possible points received. Columns 1 and 3 are restricted to individuals with a message treatment who received a functional raffle message (592 did not due to an implementation error). Columns 2 and 4 are restricted to individuals with a message treatment who took the endline survey. Omitted groups are mx-only (Column 1) and my-only (Column 2). mx & my is 1 if the individual received both sets of messages. The specification includes controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month). One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table.

Table 9: Parameter Estimates

Parameter	*Description*	*Estimate*	*Standard Error*
α	slope of marginal cost of effort	8.239	0.664
$\boldsymbol{\rho}$	**depletion**	**0.173**	**0.468**
μ_x	return to x	0.792	0.067
ϕ_x	x message attn subsidy	0.726	0.124
$\boldsymbol{\gamma}$	**overload**	**-0.047**	**0.109**
$\boldsymbol{\theta_m}$	**message diversion**	**-0.200**	**0.075**
λ	y incentive attn subsidy	3.130	0.335
ϕ_y	y message attn subsidy	1.354	0.195
σ_x	S.D. of ϵ_x, heterogeneity in return to x	1.594	0.113
σ_y	S.D. of ϵ_y, heterogeneity in return to y	1.748	0.069
$\boldsymbol{\theta_z}$	**incentive diversion**	**-0.145**	**0.184**
σ_{xy}	covariance of ϵ_x and ϵ_y	1.000	0.358
σ_{ϕ_x}	S.D. of δ_x, heterogeneity in message subsidy	1.128	0.212
σ_{ϕ_y}	S.D. of δ_y, heterogeneity in message subsidy	1.778	0.230
$\sigma_{\phi_x \phi_y}$	covariance of δ_x and δ_y	1.000	0.297

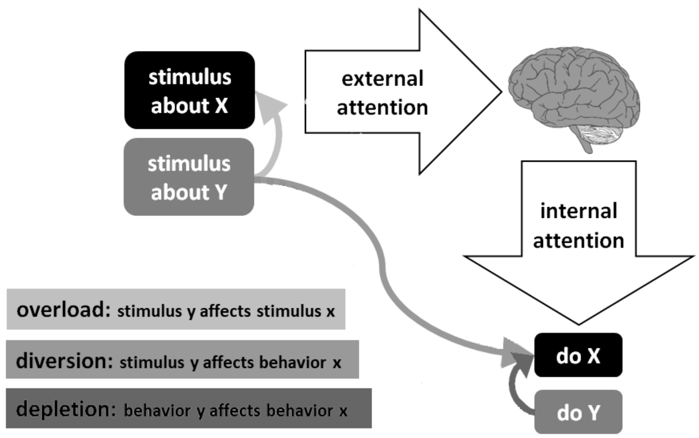
Notes: Estimates of structural parameters and their standard errors. I use classical minimum distance with 22 moments from the experiment to estimate the 15 parameters; see details in the text. Parameters linked to the three key mechanisms are written in bold.

Appendices

(For Online Publication)

A Graphical Depiction of Overload, Diversion, and Depletion

Figure A1: Overload, Depletion, and Diversion in a Taxonomy of Limited Attention



Notes: An illustration of the possible ways in which interventions could impose negative spillovers, building on a taxonomy of limited attention from Chun et al. (2011). Suppose there is an intervention, say a text message, about some behavior x, and a second intervention about some behavior y. Limits to external attention, or limited information processing, might cause the y stimulus to interfere with the x stimulus ("overload"). Limits to working or short-term memory might cause the y stimulus to divert attention toward y and away from x, reducing the likelihood of doing x ("diversion"). Finally, if the y stimulus works, causing us to do y and to exert costly cognitive effort or time, we might be subsequently less likely to do x ("depletion").

B Mathematical Proofs

B.1 Comparative Statics

$$\frac{\partial a_x^*}{\partial w_x} = \frac{-c_1^y c_3^x + c_2^x c_4^y}{c_1^x c_1^y - c_2^{x\,2}} \tag{7}$$

$$\frac{\partial a_x^*}{\partial w_y} = \frac{-c_1^y c_4^x + c_2^x c_3^y}{c_1^x c_1^y - c_2^{x\,2}} \tag{8}$$

$$\frac{\partial a_y^*}{\partial w_y} = \frac{-c_1^x c_3^y + c_2^y c_4^x}{c_1^x c_1^y - c_2^{x\,2}} \tag{9}$$

$$\frac{\partial a_y^*}{\partial w_x} = \frac{-c_1^x c_4^y + c_2^y c_3^x}{c_1^x c_1^y - c_2^{x\,2}} \tag{10}$$

$$\begin{aligned}\frac{\partial^2 a_x^*}{\partial m_x \partial m_y} &= \frac{c_1^y c_{34}^x - c_2^x c_{34}^y - c_1^y \Big((c_{11}^x \frac{\partial a_x}{\partial m_y} + c_{12}^x \frac{\partial a_y}{\partial m_y}) \frac{\partial a_x}{\partial m_x} + (c_{21}^x \frac{\partial a_x}{\partial m_y} + c_{12}^x \frac{\partial a_y}{\partial m_y}) \frac{\partial a_y}{\partial m_x} \Big)}{c_1^x c_1^y - c_2^{x\,2}} \\ &\approx \frac{c_1^y c_{34}^x - c_2^x c_{34}^y}{c_1^x c_1^y - c_2^{x\,2}}\end{aligned} \tag{11}$$

$$\begin{aligned}\frac{\partial^2 a_y^*}{\partial m_x \partial m_y} &= \frac{c_1^x c_{34}^y - c_2^x c_{34}^x - c_1^x \Big((c_{11}^y \frac{\partial a_y}{\partial m_x} + c_{12}^y \frac{\partial a_x}{\partial m_x}) \frac{\partial a_y}{\partial m_y} + (c_{21}^y \frac{\partial a_y}{\partial m_x} + c_{12}^y \frac{\partial a_x}{\partial m_x}) \frac{\partial a_x}{\partial m_y} \Big)}{c_1^x c_1^y - c_2^{x\,2}} \\ &\approx \frac{c_1^x c_{34}^y - c_2^x c_{34}^x}{c_1^x c_1^y - c_2^{x\,2}}\end{aligned} \tag{12}$$

B.2 Proof of Proposition 2

If diversion is absent and $c_4^x = c_4^y = 0$, using Equations 8 and 9, we have the following:

$$\frac{\partial a_x / \partial w_y}{\partial a_y / \partial w_y} = -\frac{c_2^x}{c_1^x} \tag{13}$$

This is an intuitive expression: it's saying that the ratio of (negative) spillover effects to main effects is equal to the ratio of depletion (c_2^x, the extent to which exerting effort on y makes it harder to exert effort on x) to effort convexity (c_1^x, the extent to which exerting effort on x makes it harder to exert more effort on x). Since neither c_2 nor c_1 vary with the intervention type, this expression remains the same for both messages ($w = m$) and incentives ($w = z$). Thus if diversion is absent, the ratio between spillover and target effects must be the same for messages about behavior y and incentives for behavior y.

B.3 Proof of Proposition 3

Re-write Equations 11 and 12 as the following, incorporating the assumption that $c_{34}^x = c_{34}^y$.

$$\frac{\partial^2 a_x^*}{\partial m_x \partial m_y} \approx \frac{-c_{34}^x (c_1^y - c_2^x)}{c_1^x c_1^y - {c_2^x}^2} \tag{14}$$

$$\frac{\partial^2 a_y^*}{\partial m_x \partial m_y} \approx \frac{-c_{34}^x (c_1^x - c_2^x)}{c_1^x c_1^y - {c_2^x}^2} \tag{15}$$

Recall that we have assumed $c_1^x > 0$ and $c_1^y > 0$. For the case where $c_2^x \leq 0$, we can quickly see from Equations 14 and 15 that $c_{34}^x = c_{34}^y > 0 \implies$ interference in both directions, and interference in either direction implies $c_{34}^x = c_{34}^y > 0$.

For the case where $c_2^x > 0$, recall that for the existence of a local maximum we have also assumed that $c_1^x c_1^y - {c_2^x}^2 > 0$ which implies that either $c_1^x > |c_2^x|$ or $c_1^y > |c_2^x|$. Then we can see that if there is interference in both directions ($\frac{\partial^2 a_x^*}{\partial m_x \partial m_y} < 0$ and $\frac{\partial^2 a_y^*}{\partial m_x \partial m_y} < 0$), then it must be true that $c_{34}^x = c_{34}^y > 0$. And if $c_{34}^x = c_{34}^y > 0$, there must be interference in at least one direction ($\frac{\partial^2 a_x^*}{\partial m_x \partial m_y} < 0$ or $\frac{\partial^2 a_y^*}{\partial m_x \partial m_y} < 0$).

C Experiment Protocol

I recruited participants using the below Facebook ad, targeting people in the U.S. age 18-65 and allowing the algorithm to train to maximize "conversions," or successful completions of Survey 1.

Figure A2: Facebook Ad Used to Recruit Participants

The first part of Survey 1 was an eligibility test in which participants had to verify six things: (1) ownership of iPhone or Android phone; (2) age 18 or over; (3) interested in working on wellness habits like daily meditation and tracking your nutrition; (4) willing to download two (free) wellness-related smartphone apps for the study; (5) comfortable potentially using a nutrition tracking app;[50] (6) have not already participated in the study. Participants then provided electronic consent, which included consenting to receive SMS messages associated with the study. The consent form also described the rewards for participation in the study: entrance into a raffle for a $20 Amazon gift card for participation in Survey 1 and app download; and entrance into a raffle for a $50 Amazon gift card for participation in Survey 2 or Survey 3. (If they completed both they were entered twice.)

The next part of Survey 1 was the app download. Participants were given two options: to either download the two apps now, or to download them after finishing the survey. Either way, they were told that they had to download both apps within 24 hours of completing the survey in order to be enrolled. They were told that if they downloaded the apps after 24 hours, they could still enroll, but they should email us to let us know. They were then shown a screen with instructions about how to download each app, which were also emailed to them upon survey

[50] We received feedback that some participants who had struggled with eating disorders or body image issues in the past were ultimately uncomfortable using the meal logging app.

completion. These instructions included a temporary assigned password, which enabled us to access their data for the duration of the study.

As described in the paper, the rest of the survey included basic demographic questions; questions on past meditation, exercise, meal logging, and sleep; questions about the full set of notifications, across devices and apps, received by the participant; and questions about the participant's perceived importance, difficulty, and "fun" of meditation, exercise, meal logging, and sleep. Upon completion of Survey 1, participants were sent an email repeating the instructions for how to download the apps, including their assigned passwords.

Every day, we verified whether new Survey 1 participants for whom 24 hours had elapsed since survey completion (plus old Survey 1 participants who emailed us) downloaded both apps. Those who did were randomized to one of the five treatments, using a script that re-randomized to ensure balance across the full sample. They were then sent an enrollment confirmation email, the key parts of which are displayed in Figure A3.[51]

Figure A3: Enrollment Confirmation Email

Welcome to the Yale Wellness and Technology Study! You successfully downloaded the apps and are officially enrolled. This email contains lots of information about the study. You can refer back to it throughout the study if you have questions.

Here is a brief summary of what it contains:

1. Your (random) assignment to messages or incentives for one or more wellness behaviors. **You were assigned: .** See below for details!
2. The **link to Survey 2**, which will expire in 24 hours. This survey is not mandatory, but takes only a few minutes, and if you participate in time, we'll enter your name into our second raffle (for a $50 Amazon gift card).
3. Survey 1 raffle results
4. A reminder of your password for [redacted]
5. Study duration and how to withdraw

Be well and let us know if you have questions.

Best,
Hannah

----

1. Your (Random) Assignment to Messages or Incentives
You have been randomly assigned to receive . [INSERT TREATMENT]

Remember, because this is an experiment, this assignment was completely random. It has nothing to do with your survey responses, or with how important we think meditation, exercise, nutrition, and sleep are. (They're all important!)

Regardless of any programs you were or were not assigned above, your ultimate use of [redacted] is entirely up to you. You are welcome but not obligated to use these apps for the study, so please use them as much or as little as you'd like. The accounts just have to stay active (with the correct email and password) for the duration of the study.

2. Link to Survey 2
Here is the link to Survey 2, which will expire in 24 hours. This survey is not mandatory, but it takes just 3-5 minutes, and if you fill it out, we'll enter your name into our second raffle for a

For example, participants received to receive messages about meditation only were told, "you have been randomly assigned to receive messages about meditation with [app], as part of our Remindful program." The goal is to make clear that the behavior we have in mind is not meditation generally, but meditation specifically with the assinged app. Below the treatment assignment, participants were given a paragraph describing the benefits of each behavior they were assigned treatment for.

[51]See Supplementary Materials for the full email.

The link to Survey 2 was included in the enrollment email. It first reminded participants of their treatment assignment, and then asked them how many days per week they "hoped" and "expected" to meditate and log their meals using the study apps. Finally, the enrollment email included Survey 1 raffle results and information about the study duration and how to withdraw. It reminded participants that at any time, they can opt out any SMS message program by replying STOP, without withdrawing from the study.

Table A1 displays the messages received by two treatment groups—meditation only and meditation and meal logging—on Day 1, in order to demonstrate the structure and timing of messages. Table A2 shows the full set of messages. (I do not reveal the names of the meditation and meal logging apps for the sake of confidentiality). The first column contains all of the messages received by any participant assigned to m_x, and the second column contains all of the messages received by any participant assigned to m_y. Each message was sent twice throughout the program (except for messages 14 and 28, which were sent just once). The first 14 rows contain the informational messages, and the second 14 rows contain the reminder/encouragement messages. (A participant assigned to, say, m_x received 2 messages per day—one informational, one reminder—over 27 days, so 54 total messages.) As mentioned in the paper, the two daily messages were sent in the morning (either 7am or 8am) and in the evening (either 7pm or 8pm). The timing of meditation vs. nutrition messages and information vs. reminder messages alternated in a balanced fashion as shown in Table A1. Messages were sent using the platform Slicktext.

The incentive treatment was described initially in the enrollment email as the following. "You will earn a green raffle ticket from eNOMerate for every day that you log at least one meal with FatSecret, and a red raffle ticket for every day that you don't. To receive a ticket, you must log a meal on the day that you ate it. Every Sunday, for the duration of the program, we will let you know via email how many tickets you've accumulated. At the end, we will pull one of your tickets, and if it's green, you will win a $10 Amazon gift certificate. So if you log your meals every day, you will definitely get the gift certificate. If you log your meals half of the time, you will get it with 50% odds. And if you never log your meals, you definitely won't get it. (This is separate from the raffles for survey completion.) The program will begin tomorrow and will last exactly 27 days."

Each Sunday, participants in the incentive treatment received an email informing them of the total green and red tickets they had accumulated. At the end of the treatment period, they were sent a final email informing them of their total tickets, and then later sent the results of the raffle. Ultimately 52% of participants won the raffle.

At the end of the treatment period, all participants received an email informing them that any treatment programs they were in would now end, but that they should keep their app accounts intact with their assigned passwords for another four weeks, when they would receive a wrap-up email from us with a link to Survey 3.

After four weeks a final email was sent, concluding the study and providing a link to Survey 3. In Survey 3, we first ask how much they meditated without the assigned apps, about the timing of

Table A1: Example Messages, Day 1

Group	Time	Msg 1	Msg 2
med only	8AM	Remember to meditate today! Try the 3-minute breathing space by Mark Williams on [meditation app]!	
	8PM	A meta-analysis in a top medical journal reviewed 47 studies and found systematic evidence that meditation reduces depression and anxiety! (Goyal et al. 2014)	
med & nut	8AM	Remember to meditate today! Try the 3-minute breathing space by Mark Williams on [meditation app]!	Logging meals can help with weight loss (Burke et al. 2011)! And people are better at meal-logging when they use apps like [meal logging app] (Wharton et al. 2014).
	8PM	A meta-analysis in a top medical journal reviewed 47 studies and found systematic evidence that meditation reduces depression and anxiety! (Goyal et al. 2014)	Take one minute to log your meals using [meal logging app] today!

Notes: This table shows the messages that were sent on Day 1, for treatment groups 2 (med only) and 4 (med & nut), as an example. (Treatment group 3 received the same nutrition messages as group 4, but without the meditation messages.) Each message program included twice-daily text messages: one simple reminder to do the behavior, and one longer message with information about some proven benefits to the behavior. Messages were sent at either 7am and 7pm or at 8am and 8pm, alternating on a daily basis. There were 14 distinct messages, and 27 days of treatment, so each message (save one) was sent twice over the course of the program. The full set of messages is shown in the Appendix in Table A2.

their meditation, and whether they felt like meditation came at the expense of any other activity. We then do the same for meal logging, with the additional question of how long it took them to log their meals each day. We then ask whether they set up any additional notifications for either behavior. Next, we ask questions about their mental health and diet. Finally, we administer an informational quiz, asking a true/false question about each informational message the participant received. At the end of Survey 3, participants were told to change their passwords for the two apps.

Table A2: Full Table of Messages

	Meditation	Nutritional Monitoring
1	Evidence from 47 studies suggests that meditation reduces depression and anxiety! (Goyal et al. 2014)	Fact: more than 102 million American adults have high cholesterol, and 35 million are at risk for heart disease as a result (CDC 2013).
2	Did you know that meditation actually changes the physical structures of the brain (Fox et al. 2014)?	Did you know that potassium helps keep your blood pressure low and your heart healthy? The CDC recommends 4700mg of potassium daily for adults age 19-50.
3	Fun fact: for people with insomnia, meditation improves nightly sleep time, and helps people fall asleep faster! (Gross et al 2011)	37.7% of Americans reported that they consume fruits less than once per day! 22.6% report the same for vegetables (CDC 2013). Make sure it's not you!
4	Aetna, a Fortune 500 company, claims that its meditation program made employees more productive, saving $3,000 per employee per year!	90% of Americans consume too much sodium (NHANES 2009-2012), which is a risk factor for heart disease! Many more foods have salt than you might expect!
5	Did you know that meditation programs combat depression almost as effectively as antidepressants? (Kuyken et al. 2008)	Over 15 years, people who consumed >25% of calories as added sugar were twice as likely to die from heart disease as those who consumed <10% (Yang et al. 2014)
6	Did you know that people can use meditation to reduce their physical pain? (Zeidan et al. 2011)	38% of U.S. adults are obese today, relative to 15% in 1980 (NHANES 2013-2014). Log your meals to keep track of your diet!
7	Fun fact: evidence suggests that meditation improves relationship satisfaction! (Sedlmeier et al. 2012)	Logging meals can help with weight loss (Burke et al. 2011)! And people are better at meal-logging when they use apps like [meal logging app] (Wharton et al. 2014).
8	Meditation programs have been shown to reduce stress levels for people with high blood pressure! (Rainforth et al. 2008)	Less than 3% of Americans meet the daily recommended fiber intake (NHANES 2003-2006). Fiber can lower cholesterol and reduce the risk of heart disease
9	Fun fact: the part of the brain responsible for memory actually looks different in people who meditate! (Fox et al. 2014)	The American Heart Association says daily consumption of added sugar should be <25g for women and <38g for men. Yet the average American consumes 82g daily.
10	Did you know that General Mills runs 7-week meditation programs for its executives? Participants say they work more productively and make better decisions.	A host of studies suggest that nutrition is the most important factor in weight management – much more important than exercise (e.g. Johns et al. 2014).
11	Meditation has so many health benefits that today, 79% of medical schools offer some element of mindfulness training (Buchholz 2015)	Are you eating enough whole grains? Find out! Whole grains reduce the risk of diabetes; refined carbohydrates actually increase the risk! (AIEssa et al. 2015)
12	Did you know that 18.1% of adults in the U.S. experience some type of anxiety disorder? Meditation has proven to help! (Goyal et al. 2014)	Moderately active women between 21-40 should be consuming 2200-2000 calories per day (and men 2600-2800). Do you? Find out by tracking meals with [meal logging app]!
13	Did you know that 35% of firms had mindfulness classes in 2017, and another 26% are considering them for the future (National Business Group on Health)?	>100 million Americans have diabetes or prediabetes (Nat'l Diabetes Stats Report 2017). Eating whole grains, and reducing sugar & trans fats, reduces the risk
14	Fun fact: meditation increases the thickness of your prefrontal cortex, the area of your brain associated with attention and self-awareness (Fox et al. 2014)	Fact: many companies are having their employees track their nutrition via smartphone apps as part of wellness programs. Jump on the bandwagon!
15	Greetings from Remindful! Try Tara Brach's Vipassana (Basic) meditation on [meditation app]!	Greetings from eNOMerate! Remember to log your meals today with [meal logging app], if you haven't already!
16	Hello from Remindful! We hope you had a great day. Try Manoj Dias' Basic Breath Meditation on [meditation app]!	Hello from eNOMerate! We hope you had a great day. Take 5 minutes to log your meals with [meal logging app]!
17	Hope you had a healthy, happy day from Remindful. You'll feel great if you end the day with some meditation! [meditation app] makes it easy.	Hope you had a healthy, happy day from eNOMerate. You'll feel great if you end the day by logging your meals! [meal logging app] makes it easy.
18	Remindful wishes you a great evening! Remember to take care of yourself, and find a few minutes to meditate with [meditation app].	eNOMerate wishes you a great evening! Remember to take care of yourself, and find a few minutes to log your meals with [meal logging app]!
19	Good evening from Remindful! You told us you were interested in meditation! So let's get on it. Try something new on [meditation app]!	Good evening from eNOMerate! You told us you were interested in monitoring your nutrition! So let's get on it. [meal logging app] makes it simple!
20	Hi from Remindful! Are you meditating daily with [meditation app]? Keep the habit up!	Hi from eNOMerate! Are you logging your meals daily with [meal logging app]? Keep the habit up!
21	Just another friendly hello, and reminder to meditate with [meditation app], from Remindful. Try the 3-minute breathing space by Mark Williams on [meditation app]!	Just another friendly hello, and reminder to log your meals with [meal logging app], from eNOMerate! ;)
22	Greetings from Remindful! Remember to meditate today with [meditation app], if you haven't already!	Greetings from eNOMerate! Remember to log your meals today with [meal logging app], if you haven't already!
23	Hello from Remindful! We hope you had a great day. Take 5 minutes to meditate with [meditation app]!	Hello from eNOMerate! We hope you had a great day. Take 5 minutes to log your meals with [meal logging app]!
24	Hope you had a healthy, happy day from Remindful. You'll feel great if you end the day with some meditation! [meditation app] makes it easy.	Hope you had a healthy, happy day from eNOMerate. You'll feel great if you end the day by logging your meals! [meal logging app] makes it easy.
25	Remindful wishes you a great evening! Remember to take care of yourself, and find a few minutes to meditate with [meditation app].	eNOMerate wishes you a great evening! Remember to take care of yourself, and find a few minutes to log your meals with [meal logging app]!
26	Good evening from Remindful! You told us you were interested in meditation! So let's get on it. Try something new on [meditation app]!	Good evening from eNOMerate! You told us you were interested in monitoring your nutrition! So let's get on it. [meal logging app] makes it simple!
27	Hi from Remindful! Are you meditating daily with [meditation app]? Keep the habit up!	Hi from eNOMerate! Are you logging your meals daily with [meal logging app]? Keep the habit up!
28	Just another friendly hello, and reminder to meditate with [meditation app], from Remindful! ;)	Just another friendly hello, and reminder to log your meals with [meal logging app], from eNOMerate! ;)

D Attrition and Survey Participation

In total 5,845 people filled out Survey 1, meaning that 66% of Survey 1 participants ultimately downloaded both apps and enrolled in the study. Of the 3,885 participants who enrolled, 40 ultimately dropped out, amounting to 1%, and resulting in a final sample of 3,845. Table A3 shows that there was no evidence of differential attrition by treatment.

Table A3: Attrition Rates by Treatment

	control	mx	my	mx & my	zy	F-test, joint sig
attrited	0.009	0.008	0.011	0.012	0.012	0.870
	0.092	0.088	0.105	0.109	0.110	

Notes: Means and standard deviations. F-test of joint significance reported in last column.

In terms of survey participation, of our 3,845 study participants, 2,891 completed Survey 2 (75.2%), and 2,145 completed Survey 3 (55.8%).

E Moment Conditions

I define ω to be $\alpha^2 - \rho^2$.

$$
\begin{aligned}
M_1 &= \mathbb{E}[a_x^*|m_x = 1, m_y = 0] - \mathbb{E}[a_x^*|m_x = 0, m_y = 0] = \frac{\alpha\phi_x - \rho\theta_m}{\omega} \\
M_2 &= \mathbb{E}[a_x^*|m_x = 0, m_y = 1] - \mathbb{E}[a_x^*|m_x = 0, m_y = 0] = \frac{\alpha\theta_m - \rho\phi_y}{\omega} \\
M_3 &= (\mathbb{E}[a_x^*|m_x = 1, m_y = 1] - \mathbb{E}[a_x^*|m_x = 0, m_y = 1]) - (\mathbb{E}[a_x^*|m_x = 1, m_y = 0] - \mathbb{E}[a_x^*|m_x = 0, m_y = 0]) \\
&= \frac{\gamma(\alpha - \rho)}{\omega} \\
M_4 &= \mathbb{E}[a_x^*|z_y = 1] - \mathbb{E}[a_x^*|z_y = 0] = \frac{\alpha\theta_z - \lambda\rho}{\omega} \\
M_5 &= \mathbb{E}[a_x^*|m_x = 0, m_y = 0, z_y = 0] = \frac{\alpha\mu_x - \rho}{\omega} \\
M_6 &= \mathbb{E}[a_y^*|m_x = 1, m_y = 0] - \mathbb{E}[a_y^*|m_x = 0, m_y = 0] = \frac{\alpha\theta_m - \rho\phi_x}{\omega} \\
M_7 &= \mathbb{E}[a_y^*|m_x = 0, m_y = 1] - \mathbb{E}[a_y^*|m_x = 0, m_y = 0] = \frac{\alpha\phi_y - \rho\theta_m}{\omega} \\
M_8 &= (\mathbb{E}[a_y^*|m_x = 1, m_y = 1] - \mathbb{E}[a_y^*|m_x = 0, m_y = 1]) - (\mathbb{E}[a_x^*|m_x = 1, m_y = 0] - \mathbb{E}[a_x^*|m_x = 0, m_y = 0]) \\
&= \frac{\gamma(\alpha - \rho)}{\omega} \\
M_9 &= \mathbb{E}[a_y^*|z_y = 1] - \mathbb{E}[a_y^*|z_y = 0] = \frac{\alpha\lambda - \rho\theta_z}{\omega} \\
M_{10} &= \mathbb{E}[a_y^*|m_x = 0, m_y = 0, z_y = 0] = \frac{\alpha - \rho\mu_x}{\omega} \\
M_{11} &= Var[a_x^*|m_x = 1, m_y = 0] - Var[a_x^*|m_x = 0, m_y = 0] = \frac{\alpha^2\sigma_{\phi_x}^2}{\omega^2} \\
M_{12} &= Var[a_x^*|m_x = 0, m_y = 1] - Var[a_x^*|m_x = 0, m_y = 0] = \frac{\rho^2\sigma_{\phi_y}^2}{\omega^2} \\
M_{13} &= (Var[a_x^*|m_x = 1, m_y = 1] - Var[a_x^*|m_x = 0, m_y = 1]) \\
&\quad - (Var[a_x^*|m_x = 1, m_y = 0] - Var[a_x^*|m_x = 0, m_y = 0]) = \frac{-2\alpha\rho\sigma_{\phi_x\phi_y}}{\omega^2} \\
M_{14} &= Var[a_x^*|m_x = 0, m_y = 0, z_y = 0] = \frac{\alpha^2\sigma_x^2 + \rho^2\sigma_y^2 - 2\alpha\rho\sigma_{xy}}{\omega^2} \\
M_{15} &= Var[a_y^*|m_x = 1, m_y = 0] - Var[a_y^*|m_x = 0, m_y = 0] = \frac{\rho^2\sigma_{\phi_x}^2}{\omega^2} \\
M_{16} &= Var[a_y^*|m_x = 0, m_y = 1] - Var[a_y^*|m_x = 0, m_y = 0] = \frac{\alpha^2\sigma_{\phi_y}^2}{\omega^2} \\
M_{17} &= (Var[a_y^*|m_x = 1, m_y = 1] - Var[a_y^*|m_x = 0, m_y = 1]) \\
&\quad - (Var[a_y^*|m_x = 1, m_y = 0] - Var[a_y^*|m_x = 0, m_y = 0]) = \frac{-2\alpha\rho\sigma_{\phi_x\phi_y}}{\omega^2} \\
M_{18} &= Var[a_y^*|m_x = 0, m_y = 0, z_y = 0] = \frac{\alpha^2\sigma_y^2 + \rho^2\sigma_x^2 - 2\alpha\rho\sigma_{xy}}{\omega^2}
\end{aligned}
$$

$$M_{19} = Cov[a_x^*, a_y^*|m_x = 1, m_y = 0] - Cov[a_x^*, a_y^*|m_x = 0, m_y = 0] = \frac{-\alpha\rho\sigma_{\phi_x}^2}{\omega^2}$$

$$M_{20} = Cov[a_x^*, a_y^*|m_x = 0, m_y = 1] - Cov[a_x^*, a_y^*|m_x = 0, m_y = 0] = \frac{-\alpha\rho\sigma_{\phi_y}^2}{\omega^2}$$

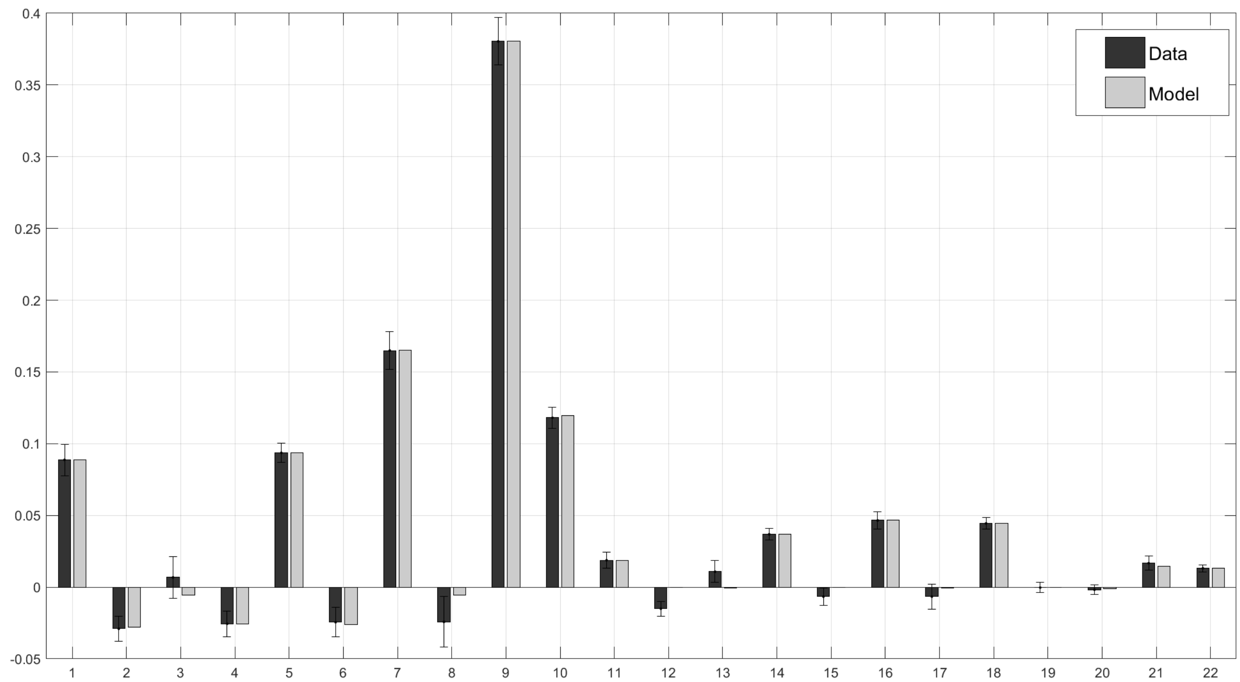
$$M_{21} = (Cov[a_x^*, a_y^*|m_x = 1, m_y = 1] - Cov[a_x^*, a_y^*|m_x = 0, m_y = 1])$$

$$- (Cov[a_x^*, a_y^*|m_x = 1, m_y = 0] - Cov[a_x^*, a_y^*|m_x = 0, m_y = 0]) = \frac{(\alpha^2 + \rho^2)\sigma_{\phi_x\phi_y}}{\omega^2}$$

$$M_{22} = Var[a_x^*, a_y^*|m_x = 0, m_y = 0, z_y = 0] = \frac{-\rho(\alpha\sigma_x^2 + \alpha\sigma_y^2) + (\alpha\alpha + \rho^2)\sigma_{xy}}{\omega^2}$$

F Estimation Robustness Checks and Sensitivity

Figure A4: Model Fit



Notes: Comparison of actual moments to predicted moments.

In Figure A4 I illustrate the model fit by comparing the moments in the data to those predicted by the model. In Table A4 I show parameter estimates using the identity matrix as the weighing matrix, instead of the diagonal of the inverse of the variance-covariance matrix as in the benchmark specification. In Figure A5 I plot the sensitivity matrix, as defined by Andrews et al. (2017), showing only the four key parameters. Each element of the matrix represents how a one percentage point increase in each moment affects each estimated parameter. We can see that ρ is most sensitive to the treatment effects of messages on the covariance, as well as to the main target effects, spillover effects, and interaction effects. As we would expect, θ_m and θ_z are sensitive to the same moments as ρ, but mostly in the opposite direction, since spillovers must be due to either diversion or depletion. (The bars go in the same direction because depletion is captured by $\rho > 0$ but diversion is captured by $\theta < 0$.) The only major exception is the effect of m_y on x, which pushes both θ_m and ρ toward zero. This is intuitive: the closer this spillover gets to zero, the smaller are both estimates of diversion and depletion. The effect of m_x on y pushes them in opposite directions (θ_m toward zero; ρ away from zero) because this moment is being used to separate ρ from θ_m: the nearer to zero is this spillover, the smaller is the spillover-target ratio for x, bringing it closer to the spillover-target ratio for y and thus making ρ larger. As would be expected. γ is most

Table A4: Estimates using Identity Matrix

Parameter	*Description*	*Estimate*	*Standard Error*
α	slope of marginal cost of effort	8.301	0.574
$\boldsymbol{\rho}$	**depletion**	**0.159**	**0.711**
μ_x	return to x	0.796	0.080
ϕ_x	x message attn subsidy	0.732	0.110
$\boldsymbol{\gamma}$	**overload**	**-0.073**	**0.113**
$\boldsymbol{\theta_m}$	**message diversion**	**-0.202**	**0.133**
λ	y incentive attn subsidy	3.154	0.289
ϕ_y	y message attn subsidy	1.365	0.163
σ_x	S.D. of ϵ_x, heterogeneity in return to x	1.605	0.101
σ_y	S.D. of ϵ_y, heterogeneity in return to y	1.760	0.078
$\boldsymbol{\theta_z}$	**incentive diversion**	**-0.152**	**0.281**
σ_{xy}	covariance of ϵ_x and ϵ_y	0.992	0.439
σ_{ϕ_x}	S.D. of δ_x, heterogeneity in message subsidy	1.136	0.193
σ_{ϕ_y}	S.D. of δ_y, heterogeneity in message subsidy	1.791	0.185
$\sigma_{\phi_x \phi_y}$	covariance of δ_x and δ_y	1.000	0.362

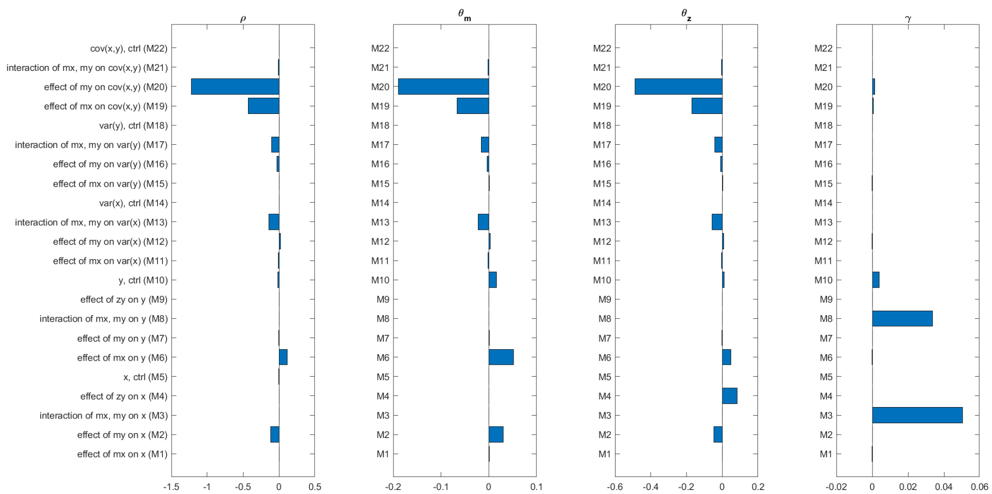
Notes: Parameter estimates using the identity matrix as the weighing matrix, instead of the diagonal of the inverse of the variance-covariance matrix as in the benchmark specification.

Table A5: Estimates with θ_x and θ_y

Parameter	*Description*	*Estimate*	*Standard Error*
α	slope of marginal cost of attention	8.353	0.602
$\boldsymbol{\rho}$	**depletion**	**0.036**	**0.265**
μ_x	return to x	0.786	0.063
ϕ_x	x message attn subsidy	0.740	0.117
$\boldsymbol{\gamma}$	**overload**	**-0.047**	**0.109**
$\boldsymbol{\theta_x}$	**diversion generated by x interventions**	**-0.201**	**0.078**
λ	y incentive attn subsidy	3.176	0.320
ϕ_y	y message attn subsidy	1.377	0.183
σ_x	S.D. of ϵ_x, heterogeneity in return to x	1.606	0.118
σ_y	S.D. of ϵ_y, heterogeneity in return to y	1.764	0.061
$\boldsymbol{\theta_y}$	**diversion generated by y interventions**	**-0.219**	**0.085**
σ_{xy}	covariance of ϵ_x and ϵ_y	0.943	0.261
σ_{ϕ_x}	S.D. of δ_x, heterogeneity in message subsidy	1.144	0.207
σ_{ϕ_y}	S.D. of δ_y, heterogeneity in message subsidy	1.803	0.220
$\sigma_{\phi_x \phi_y}$	covariance of δ_x and δ_y	1.000	0.335

Notes: Parameter estimates, replacing parameters θ_m and θ_z with θ_x and θ_y (allowing diversion to vary across behaviors but not across interventions).

Figure A5: Sensitivity of Estimates to Moments



Matrix depicting sensitivity of parameter estimates to moments, as defined by Andrews et al. (2017), showing only the four key parameters. Each element of the matrix represents how a one percentage point increase in each moment affects each estimated parameter.

sensitive to the message interaction moments, though it is sensitive to all of the other moments through ρ and α.

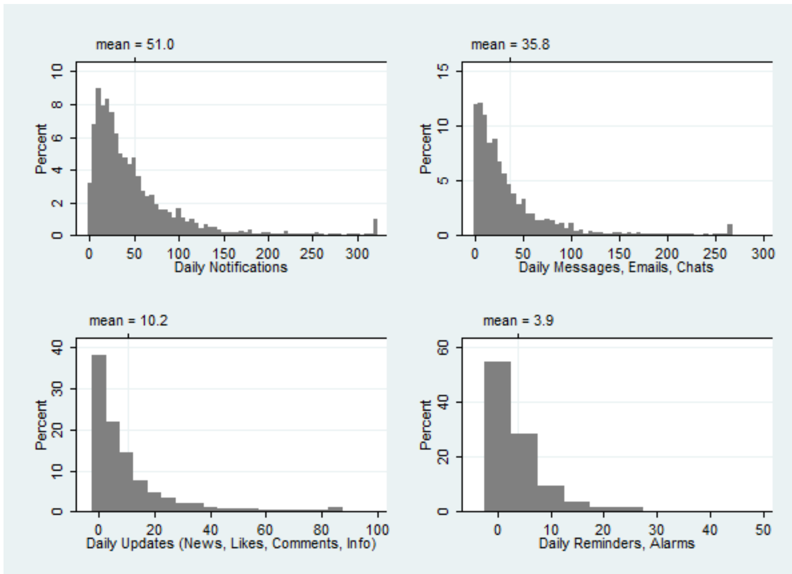
G What would need to be the elasticity of substitution for the time constraint to explain spillovers?

Suppose agents have CES utility, $u(h_x, h_y) = (\alpha h_x^\rho + (1-\alpha)h_y^\rho)^{(1/\rho)}$, where h_x and h_y are daily hours spent on meditation and meal logging, respectively. It can be easily shown that the cross-price elasticity of meditation with respect to meal logging ϵ_{xy} is equal to $(\sigma - 1) * s_y$, where $s_y = \frac{h_i p_i}{Y}$, the share of income spent on y, and $\sigma = \frac{1}{1-\rho}$, the elasticity of substitution. In our case, $s_y = \frac{h_i}{24}$, if I let p_i represent the value of 1 hour of time.

I can use my estimates to approximate a lower bound for ϵ_{xy}. Table A10 shows that daily average minutes meditated fell from 1.8 to 0.98, a 45% decrease, in the presence of incentives for meal logging. The incentives had an expected value of $0.37 per successful day (with at least one meal logged), so I assume the price of meal logging fell by this amount. I approximate the price of meal logging to be $20 * (3.78/60) = 1.26$, assuming that the average hourly wage for a college graduate is $20, and that the minimum necessary time spent on meal logging was 3.78 minutes (the time required to log one meal as reported in the final survey, which is a lower bound for the time spent on daily meal logging). Together, this implies that incentives for meal logging reduced the price of meal logging by 29.4%, and that the cross price elasticity, ϵ_{xy}, is equal to 1.53. To calculate h_y, I use the time required to log all of one's meals as reported in the final survey, 11.82 minutes (an upper bound for the time spent on daily meal logging) and multiply it by the average daily rate of logging at least one meal across all groups, which was 0.14. Combining everything, I find that in order to explain the spillovers we see, the elasticity of substitution would need to be at least 1331.

H Additional Data & Analysis

Figure A6: Daily Notifications (after winsorizing at 99%)



Notes: The distribution of daily notifications, as self-reported in the baseline survey. Participants were asked to list all apps that send notifications across all devices, and then to estimate daily notifications for each app. The top-left plot shows total notifications, and the subsequent plots break notifications down by type.

Table A6: Expectations and Hopes: Meditation, Meal Logging, Exercise, and Sleep

Panel A: Expectations

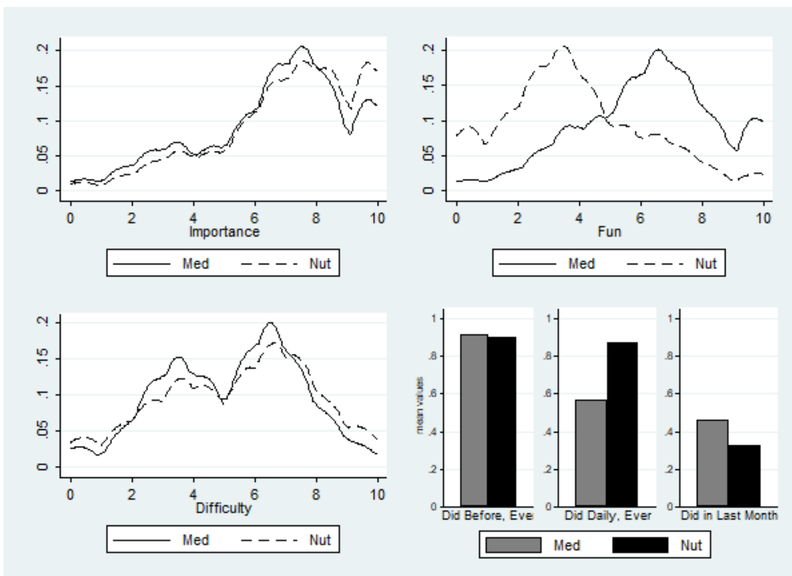
	Expected to Meditate (X) (1)	Expected to Meal Log (Y) (2)	Expected to Exercise (3)	Expected to Sleep (4)
mx	0.113***	-0.017	0.022	-0.008
	(0.012)	(0.017)	(0.014)	(0.015)
my	0.017	0.122***	0.001	0.003
	(0.012)	(0.014)	(0.013)	(0.014)
mx X my	-0.027	0.037	-0.020	0.000
	(0.017)	(0.021)	(0.019)	(0.021)
zy	0.021	0.213***	0.012	0.000
	(0.013)	(0.015)	(0.014)	(0.015)
mx + mxmy	0.085	0.019	0.002	-0.008
	(0.012)	(0.013)	(0.013)	(0.014)
my + mxmy	-0.010	0.159	-0.019	0.003
	(0.012)	(0.015)	(0.014)	(0.015)
Ctrl Mean	0.384	0.519	0.475	0.594
Ctrl Mean S.E.	(0.009)	(0.012)	(0.010)	(0.010)
Obs	2891	2891	2891	2891

Panel B: Hopes

	Hoped to Meditate (X) (1)	Hoped to Meal Log (Y) (2)	Hoped to Exercise (3)	Hoped to Sleep (4)
mx	0.126***	-0.016	0.025	-0.001
	(0.013)	(0.018)	(0.012)	(0.010)
my	0.038**	0.102***	0.001	0.004
	(0.014)	(0.015)	(0.012)	(0.010)
mx X my	-0.048**	0.034	-0.028	-0.009
	(0.019)	(0.022)	(0.017)	(0.014)
zy	0.028	0.166***	0.004	0.009
	(0.015)	(0.014)	(0.013)	(0.010)
mx + mxmy	0.078	0.018	-0.003	-0.010
	(0.014)	(0.012)	(0.012)	(0.010)
my + mxmy	-0.011	0.136	-0.027	-0.005
	(0.012)	(0.016)	(0.012)	(0.010)
Ctrl Mean	0.643	0.779	0.681	0.906
Ctrl Mean S.E.	(0.011)	(0.013)	(0.009)	(0.007)
Obs	2891	2891	2891	2891

Notes: Expected rates of behavior at baseline over treatment period (Panel A) and hoped-for rates of behavior at baseline over treatment period (Panel B). Regressions include controls for the five baseline variables on which re-randomization was based. One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Figure A7: Preferences and Experience, Meditation & Meal Logging



Notes: The distribution of baseline responses to questions about self-reported importance, fun, and difficulty of each behavior, on a scale from 1 to 10. The most notable difference between the two behaviors is that participants believe that meditation will be more "fun" than meal logging. In the bottom-right plot I depict the self-reported experience with each behavior. The first comparison shows the fraction of participants who ever did the behavior before, the second shows the fraction of participants who ever did the behavior *daily* before, and the third shows the fraction of participants who did the behavior in the last month.

Table A7: Reduced Form Results Reported as Treatment Effects

	Binary Outcomes		*Continuous Outcomes*	
	Meditated (0/1) (1)	Logged Meal (0/1) (2)	Min. Meditated (3)	Cal. Logged (4)
mx only	0.088***	-0.024**	1.298***	-35.601**
	(0.011)	(0.010)	(0.313)	(15.110)
my only	-0.028***	0.166***	-0.863***	212.193***
	(0.008)	(0.013)	(0.245)	(19.473)
mx & my	0.066***	0.116***	0.867***	150.646***
	(0.010)	(0.012)	(0.296)	(18.291)
zy	-0.025***	0.381***	-0.868***	350.546***
	(0.009)	(0.016)	(0.247)	(22.258)
mx - mx & my	0.022	-0.140	0.430	-186.248
	(0.011)	(0.012)	(0.319)	(17.736)
my - mx & my	-0.094	0.050	-1.730	61.546
	(0.009)	(0.014)	(0.252)	(21.582)
Ctrl Mean	0.094	0.118	2.090	153.352
Ctrl SD	0.291	0.323	8.705	473.961
Obs	102905	102905	102905	102905

Notes: OLS regressions of outcomes (binary in columns 1 and 2, continuous in columns 3 and 4) on treatments. mx only is 1 if the individual received only meditation messages and 0 otherwise; my only is 1 if the individual received only meal logging messages and 0 otherwise; mx & my is 1 if the individual received both meditation and meal logging messages and 0 otherwise; zy is 1 if the individual received only meal logging incentives and 0 otherwise. Specifications include controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1\%, 5\%, and 10\% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table A8: Reduced Form Results, Including Use of Other Apps

	Treatment Period		*Post-Treatment Period*	
	Meditated (x) (1)	Logged Meal (y) (2)	Meditated (x) (3)	Logged Meal (y) (4)
mx	0.242***	-0.032	0.028*	-0.005
	(0.032)	(0.037)	(0.014)	(0.010)
my	-0.027	0.397***	-0.014	0.059***
	(0.029)	(0.036)	(0.011)	(0.013)
mx X my	0.023	-0.084	0.015	-0.013
	(0.046)	(0.051)	(0.019)	(0.018)
zy	-0.011	0.450***	-0.029**	0.172***
	(0.030)	(0.036)	(0.011)	(0.018)
mx + mxmy	0.265	-0.116	0.043	-0.018
	(0.034)	(0.036)	(0.013)	(0.015)
my + mxmy	-0.003	0.313	0.001	0.046
	(0.036)	(0.037)	(0.015)	(0.012)
Ctrl Mean	0.346	0.557	0.064	0.047
Ctrl SD	0.460	0.574	0.245	0.211
Obs	2131	2119	3805	3805

Notes: OLS regressions of an individual-level outcome variable that takes into account the use of other meditation and meal logging apps on treatments. At the final survey, we ask participants how many days they did the behaviors using other apps during the treatment and post-treatment period. I inflate mean meditation and meal logging rates for the duration of the period according to the number of days in which other apps were reported to be used. The specification includes controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month). One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table A9: Reduced Form Results, Intensive Margin

	Treatment Period	
	Min. Meditated, Cond'l on Meditating (3)	Cal. Logged, Cond'l on Logging (4)
mx	-3.619**	-42.349
	(1.457)	(41.762)
my	-3.663*	-1.658
	(1.756)	(35.262)
mx X my	3.557	47.483
	(2.065)	(49.633)
zy	-4.756**	-273.184***
	(1.714)	(36.700)
mx + mxmy	-0.062	5.134
	(1.483)	(26.878)
my + mxmy	-0.105	45.824
	(1.116)	(35.017)
Ctrl Mean	22.322	1299.512
Ctrl SD	18.919	643.607
Obs	11848	24309

Notes: OLS regressions of continuous outcomes, conditional on doing the behavior, on treatments. mx is 1 if the individual received meditation messages and 0 otherwise; my is 1 if the individual received meal logging messages and 0 otherwise; mx X my is the interaction between mx and my, zy is 1 if the individual received meal logging incentives. Specifications include controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1\%, 5\%, and 10\% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table A10: Treatment Effects on Health Outcomes

	Avg Daily Min Meditated (1)	Standardized PHQ4 Score (2)	Standardized M. Health Score (3)	Fraction Weight Goal Achieved (4)	Standardized Diet Score (5)
mx	0.778**	-0.010	0.079	-0.060	0.103
	(0.288)	(0.065)	(0.070)	(0.057)	(0.065)
my	-0.861***	-0.007	0.100	0.012	0.224***
	(0.225)	(0.066)	(0.067)	(0.065)	(0.070)
mx X my	0.528	-0.069	0.006	0.003	-0.120
	(0.364)	(0.092)	(0.101)	(0.078)	(0.097)
zy	-0.818***	0.006	-0.058	-0.019	0.302***
	(0.234)	(0.068)	(0.066)	(0.055)	(0.072)
mx + mxmy	1.306	-0.079	0.085	-0.057	-0.017
	(0.222)	(0.066)	(0.072)	(0.054)	(0.072)
my + mxmy	-0.333	-0.076	0.106	0.015	0.104
	(0.286)	(0.065)	(0.075)	(0.037)	(0.068)
Ctrl Mean	1.800	0.000	0.000	0.132	0.000
Ctrl Mean S.D.	5.633	1.000	1.000	0.964	1.000
Obs	3826	2145	2141	1651	2142

Notes: Health outcomes, including (1) average daily minutes meditated; (2) standardized score from the PHQ4, a four-item anxiety and depression questionnaire (specifically, respondents are diagnosed as having levels of depression/anxiety that are “normal,” “mild,” “moderate,” or “severe” according to standard score cut-offs; I then score these diagnoses as 0, 1, 2, or 3, respectively, and then standardize relative to the control group, where higher z-scores represent lower mental health); (3) standardized response to ”How would you describe your mental health now, relative to before you started the study?”; (4) fraction of weight goal achieved (self-reported), and (5) standardized response to ”How would you describe your diet now, relative to before you started the study?” Regressions include controls for the five baseline variables on which re-randomization was based. One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

I Heterogeneity

I look at two potential sources of heterogeneity in spillover effects. The first is by baseline notifications. Since the vast majority (92%) of notifications received by our sample participants are messages or updates, and not associated with a particular behavior, I simply modify the model to include a new type of message, m_w, but do not allow for any action associated with w (i.e. a_w). I make the same assumptions as before. The components of the cost function become: $f(a_x, a_y) = \frac{1}{2}\alpha(a_x^2 + a_y^2) + \rho a_x a_y$ and $s^x(m_x, z_x, m_y, z_y, m_w, z_w) = \nu_x m_x + \gamma m_x m_y + \gamma m_x m_w + \lambda z_x + \theta_m m_y + \theta_m m_w + \theta_z z_y$. I find the effect of m_w on spillovers to be:

$$\begin{aligned}(\mathbb{E}[a_x^*|m_x = 0, m_y = 1, m_w = 1] - \mathbb{E}[a_x^*|m_x = 0, m_y = 0, m_w = 1])\\ - (\mathbb{E}[a_x^*|m_x = 0, m_y = 1, m_w = 0] - \mathbb{E}[a_x^*|m_x = 0, m_y = 0, m_w = 0]) = \frac{-\rho\gamma}{\alpha^2 - \rho^2}\end{aligned}$$

In the presence of both depletion ($\rho > 0$) and overload ($\gamma < 0$), notifications are predicted to push spillovers of m_y on x toward zero. The intuition is that distracting messages m_w interfere with m_y due to overload, causing them to generate less depletion, and thus less of a spillover, than they otherwise would. Since I found no strong evidence of either depletion or overload, I should not expect to see strong heterogeneous effects by baseline notifications. I do not attempt to predict the continuous effect of additional notifications, which would require many more assumptions.

The most accurate test of the above prediction would be to look at people who have at least one other daily notification versus people who do not. Unfortunately only 12 participants have fewer than one other daily notification, and even cutting the data in half does not provide satisfactory power to detect small effects. Instead I opt for the test with the highest power, interacting each treatment with whether participants have notifications that are above or below the median. Table A11 shows the results. As expected, there is limited evidence of heterogeneity, though for all of the above reasons this is by no means a conclusive test.

I do not attempt to use the model to predict heterogeneity by baseline experience, as it is not clear how experience should enter the model. It likely reflects both preferences, which would enter through u, as well as accumulated habits, which would likely affect the cost of attention. Instead, I try to run the simplest possible test of heterogeneity by baseline experience, in the hope that it might provide insight for future models.

I construct an experience score in which the participant gets 1 point if he/she has ever done the behavior, another point if he/she has attempted to do it daily before, and another point if he/she has done it in the last month, for a minimum score of zero and a maximum of three. Table A12 shows the results by whether participations are above or below the median in their experience with the outcome behavior in question. (In Column 1, experience represents meditation experience; in Column 2, experience represents meal logging experience.) I find no evidence of heterogeneity by experience, but again, the data is under-powered to detect small effects.

Table A11: Heterogeneous Treatment Effects by Baseline Notifications

	Meditated (x)	Logged Meal (y)
	(1)	(2)
mx	0.099***	-0.032*
	(0.016)	(0.015)
my	-0.043***	0.167***
	(0.012)	(0.019)
mx X my	0.012	-0.005
	(0.021)	(0.026)
zy	-0.028*	0.397***
	(0.013)	(0.024)
highnotif	-0.013	-0.012
	(0.014)	(0.015)
mx X highnotif	-0.024	0.017
	(0.021)	(0.021)
my X highnotif	0.031	-0.002
	(0.017)	(0.026)
mx X my X highnotif	-0.010	-0.041
	(0.028)	(0.035)
zy X highnotif	0.005	-0.031
	(0.018)	(0.033)
mx + mxmy	0.111	-0.038
	(0.013)	(0.021)
my + mxmy	-0.032	0.162
	(0.017)	(0.018)
(mx + mxmy) X highnotif	-0.030	-0.020
	(0.020)	(0.030)
(my + mxmy) X highnotif	0.020	-0.040
	(0.020)	(0.020)
Ctrl Mean	0.094	0.118
Ctrl SD	0.291	0.323
Obs	102905	102905

Notes: OLS regressions of treatment-period behaviors on treatments and interactions with a binary measure of whether daily notifications are above or below the median. Includes controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

Table A12: Heterogeneous Effects by Baseline Experience in Outcome Behavior

	Meditated (x) (1)	Logged Meal (y) (2)
mx	0.093***	-0.021
	(0.021)	(0.026)
my	-0.002	0.133***
	(0.014)	(0.030)
mx X my	-0.045	0.015
	(0.027)	(0.044)
zy	-0.009	0.358***
	(0.015)	(0.045)
experience	-0.008	0.004
	(0.011)	(0.028)
mx X experience	-0.003	-0.001
	(0.011)	(0.012)
my X experience	-0.013	0.016
	(0.008)	(0.014)
mx X my X experience	0.026	-0.020
	(0.015)	(0.020)
zy X experience	-0.008	0.011
	(0.009)	(0.020)
mx + mxmy	0.048	-0.006
	(0.017)	(0.036)
my + mxmy	-0.047	0.148
	(0.023)	(0.033)
(mx + mxmy) X experience	0.020	-0.020
	(0.010)	(0.020)
(my + mxmy) X experience	0.010	0.000
	(0.010)	(0.010)
Ctrl Mean	0.094	0.118
Ctrl SD	0.291	0.323
Obs	102905	102905

Notes: OLS regressions of treatment-period behaviors on treatments and interactions with a binary measure of whether baseline experience in the outcome behavior was above or below the median. The experience measure went from 0 to 3, where participants earned 1 point for having ever done it before, 1 point for having done it daily before, and 1 point for having done it in the last month. Includes controls for the five baseline variables on which re-randomization was based (female, college, daily notifications, whether individual meditated in last month, whether individual logged meal in last month) as well as day fixed effects. Standard errors clustered at individual level. One, two, and three stars indicate q-values of 1%, 5%, and 10% respectively; q-values calculated according to the Benjamini Hochberg step-down procedure, considering all tests in the table (but excluding linear combinations of coefficients).

J Deviations from Pre-Analysis Plan

This study was registered at the AEA RCT Registry under the title ”Nudges in Equilibrium” with RCT ID AEARCTR-0002435. In this section I describe any differences between the final paper and the pre-analysis plan.

Importantly, the main experiment design and sample size did not change substantially between the pre-analysis plan and the experiment. Slight differences in sample size across treatment groups are due to the fact that we randomized within cohorts using fixed proportions, and did not have full control over the total numbers. The slight rise in the total sample is also due to being unable to exactly control the size of the final cohort.

In terms of the analysis, the reduced form specifications are the same. One important difference is that ultimately we used meditation and meal logging *with* the assigned apps as the outcome in our main specification, rather than incorporating self-reports of meditation and meal-logging with other apps, as planned. The reason for this is twofold. First, it was actually a mistake to plan to incorporate self-reports, because the behavior we promoted in both message and incentive treatment was the behavior using the specified app, not the behavior generally. Second, ultimately only 56% of participants completed Survey 3—much less than hoped—so incorporating self-reports from this survey reduced our power significantly. Table A8 shows the results of the specification stated in the pre-analysis plan. The key coefficients of interest are not substantially different, but there is insufficient power to draw the same conclusions.

I do not include in the paper all of the sub-group analyses as described in the pre-analysis plan, since they are generally insufficiently powered. I also changed the measurement tool for mental health, substituting the PHQ4 for the General Well-Being Schedule, since feedback from the first participants suggested that the 18-item questionnaire was too time-consuming, and was reducing the likelihood of completion.

Finally, the model changed in several ways since the pre-analysis plan. The most important change was to ultimately include three mechanisms (overload, diversion, and depletion), instead of just two (overload and depletion, originally called limited external and internal attention). I made this change because after getting the data, I realized that the original model simply did not fit because it did not include diversion, which is what I ultimately find to be the main driver of spillovers. As a result of this change, many smaller, subsequent changes had to be made to the model as well, which explains the rest of the differences.