On the Failure to Seek Beneficial Information: The Problem with Inconspicuous Incentives

Leslie K. John
Hayley Blunden
Katherine Milkman
Luca Foschini
Francesca Gino
Bradford Tuckfield

Working Paper 16-090

On the Failure to Seek Beneficial Information: The Problem with Inconspicuous Incentives

Leslie K. John
Harvard Business School

Luca Foschini
Evidation Health

Hayley Blunden
Harvard Business School

Francesca Gino
Harvard Business School

Katherine Milkman
University of Pennsylvania

Bradford Tuckfield
American Express

On the Failure to Seek Beneficial Information: The Problem with Inconspicuous Incentives

Leslie K. John
Harvard University

Hayley Blunden
Harvard University

Katherine Milkman
University of Pennsylvania

Luca Foschini
Evidation Health

Francesca Gino
Harvard University

Bradford Tuckfield
American Express

Abstract: Managers and policymakers regularly rely on incentives to encourage valued behaviors. While often successful, there are also notable and surprising examples of their ineffectiveness. Why? Perhaps they are not sufficiently conspicuous. In support of this account, in a large-scale field experiment and laboratory study, we show that even when incentives are transparently provided and easily trackable in real time, failing to make them conspicuous renders incentives ineffectual at shifting behavior. Further, we show that inconspicuous incentives are ineffective in part because people fail to seek information on how to improve their outcomes *despite* being informed that such information exists and *even when* it can be obtained at the mere click of button. Finally, we show that people fail to appreciate potential beneficiaries' apparent lack of interest in obtaining incentive information, a result suggesting that purveyors of incentive programs may under-invest in promoting them.

Introduction

Managers and governments often rely on incentive programs to change people's behavior, encouraging valued actions such as charitable giving, home ownership, smoking cessation, vaccination, exercise, energy conservation, and savings. Although these programs are often highly successful (Bronchetti, Huffman, & Magenheim, 2015; Charness & Gneezy, 2009; Imas, 2014; Royer, Stehr, & Sydnor, 2015; Sindelar, 2008; Volpp et al., 2008; Volpp et al., 2009), there are also notable and surprising examples of their ineffectiveness (e.g., Gneezy, Meier, & Rey-Biel, 2011). For instance, the Federal government offered the Hope, Lifetime Learning, and American Opportunity Tax Credits to subsidize spending on higher education, and these incentive programs had a negligible effect on college enrollments (Bulman & Hoxby, 2015; Long, 2004). One explanation could be that the college enrollment decision is highly inelastic, but enrollment decisions have been shown to respond dramatically when tax professionals simply help families fill out financial aid forms (Bettinger, Long, Oreopoulos, & Sanbonmatsu, 2012).

Another puzzling failure of incentives occurred in a California program that offered discounts on electricity for consumers who reduced their energy use by 20%. This program produced no measurable response (Ito, 2015), but mailings merely letting consumers know how their energy usage compared to that of their neighbors reliably and meaningfully reduced energy consumption (Allcott, 2011), as did messages associating pollution from energy usage with health and environmental problems (Asensio & Delmas, 2015). These examples and many others present a puzzle: why don't incentives to change highly elastic decisions always produce the desired response? After all, standard economic theory assumes that people will respond to

positive incentives by increasing their engagement in the incentivized behavior (Mas-Colell, Whinston, & Green, 1995).

Recent work has called this assumption into question, suggesting a number of key boundary conditions ranging from paying so little that people are insulted (Gneezy & Rustichini, 2000b; Heyman & Ariely, 2004), to crowding out intrinsic motivation (Deci, 1971; Gneezy & Rustichini 2000). Other work has highlighted the important roles of pre-commitment (Baca-Motes, Brown, Gneezy, Keenan, & Nelson, 2012; Royer, Stehr, & Sydnor, 2015; Schwartz et al., 2014) and perceived self-efficacy (Bandura, Adams, & Beyer, 1977) in facilitating behavior change. In this paper, we highlight another important boundary condition to incentives' effectiveness. Specifically, we show that people can be inattentive to information about incentives, and we propose that this inattentiveness can help explain why incentives designed to change highly elastic decisions may often fail. Further, we propose and show that people are bad at predicting one another's inattentiveness, which leads them to under-invest in making incentives conspicuous. One implication of this account is that making incentives more salient – i.e., conspicuous – could increase their impact. A small but growing body of recent work points towards the value of salience in consumer decision making (Chetty, Looney, & Kroft, 2009; Tiefenbeck et al., 2018; Luca & Smith, 2013). We build on this research to assess how conspicuousness influences the effect of incentive information, showing in the field, and more explicitly in two experiments, that, contrary to standard economic theory and human intuition, unless information on the available incentives is transparently conspicuous, it may have absolutely no effect.

Specifically, we conducted a large-scale randomized field experiment of a highly consequential behavior: physical activity (425,000 Americans die prematurely each year due to

physical inactivity and diet; Mokdad, Marks, Stroup, & Gerberding, 2004). For two weeks, we offered users of a pedometer app a 40x increase in the incentives they were typically offered for meeting daily step targets, and we provided them with constant, real time access to information about the rewards they were accruing. For half of participants, we implemented a shoestring email campaign intended to increase the conspicuousness of the incentives. Our results indicated that the 40x increase in incentives was, in and of itself, insufficient to change behavior: Physical activity levels were no higher among participants who were offered inconspicuous incentives than among participants in a holdout comparison group that received the standard (40x smaller) incentives. Only when our incentive increase was made conspicuous, via a low-touch email campaign, did it boost exercise (increasing daily step counts by 3-7%, depending on the estimation model, for every day of the two week program and the two weeks that followed). Simply put, our field experiment shows that the conspicuousness of incentives is critical to their impact.

We then replicate and extend this result in two lab experiments. In the first study, we replicate the findings from our field experiment in the laboratory using a more conservative manipulation of incentive conspicuousness. In the second study, we assess people's motivation to seek information about incentives that would make them better off (i.e., instructions on how to earn more money). We show that approximately two thirds of study participants are unwilling to so much as click on a link to learn about incentives – i.e., they are unwilling to engage in a virtually costless behavior to find out how they could earn more money. As a result, these participants earn less money than they would have, had they been exposed to this incentive information. Finally, we show that people fail to appreciate others' lack of interest in learning about incentives. This result suggests that purveyors of incentive programs may under-appreciate

the need to make incentives conspicuous if they hope to achieve their intended impact. Taken together, our research highlights that when incentive programs are not made sufficiently conspicuous, even the best-intended, well-structured programs are likely to have far less impact than they could and should.

The remainder of this paper is organized as follows. First we review the literature on the effectiveness of incentives for behavior change, on inattentiveness to incentives and on people's general naiveté about their and others' failures to respond optimally to their environments. We do so with an emphasis on developing our hypotheses and motivating our studies. Then, we describe our field experiment in which increased financial incentives to exercise were only effective when communicated conspicuously. Next, we present two lab experiments in which we replicate and extend this result. We then discuss the implications of our results for the numerous incentive programs designed by policymakers to affect behavior change, many of which offer inconspicuous incentives and are thus likely less effectual than they could be. Finally, we conclude and discuss directions for future research.

Conceptual development

There is extensive past research exploring how financial incentives can be deployed to support behavior change (for a review, see Kamenica, 2012).[1] This literature generally supports the conclusion that incentives "work" – i.e., they induce behavior change. Consistent with this notion, researchers have identified neurological markers of incentives (McClure, Daw, & Montague, 2003). However, incentives can backfire – i.e., dissuade people from engaging in the

---

[1] By "incentives," we giving people cash for engaging in a target behavior (as opposed to say, offering taxes or subsidies for doing so).

target behavior – when they crowd-out intrinsic motives (Deci, 1971; Gneezy & Rustichini, 2000a), or when they are perceived as either inappropriately large (Baumeister, 1984; Ariely, Gneezy, Loewenstein, & Mazar, 2009) or inappropriately small (Gneezy & Rustichini, 2000b; Heyman & Ariely, 2004). Of particular relevance to the present investigation, theory suggests that incentives may be particularly valuable when it comes to encouraging behaviors that people tend to (perpetually) put off (Frederick, Loewenstein, O'Donoghue, 2011; Laibson, 2011; Loewenstein, & Thaler, 1989; O'Donoghue, & Rabin, 2011). That is, behaviors, such as exercise, whereby the "costs" (e.g., physical exertion) are experienced in the present, and the benefits (e.g., health, more attractive physical appearance), are delayed, and typically difficult to attribute to any given (exercise) episode. In such cases, financial incentives may hold promise in rectifying this imbalance, giving individuals an immediate benefit to offset the immediate costs of engaging in the desired behavior (Reimann, MacInnis, & Bechara, 2013).

Consistent with this account, a meta-analysis on incentives to engage in healthy behaviors, such as getting a flu shot or quitting smoking, concluded financial incentives tend to be highly effective at inducing behavior change (Giles et al., 2014) – at least in the short run. Relatedly, additional studies have shown that financial incentives can facilitate weight loss (Finkelstein, Linnan, Tate, & Birken, 2007; John et al., 2011; Volpp et al., 2008), though their impact varies depending on the incentive's size and schedule (Jeffery, 2012). Similarly, a meta-analysis that looked exclusively at exercise concluded that financial incentives can be effective – especially for those who were sedentary to begin with (Mitchell et al., 2013). In all of this research, in which financial incentives were found to be effective, however, the financial incentives offered to people for behavior change were highly conspicuous. For example, in Volpp et al. (2008), a researcher informed participants of the incentives on offer via a one-on-

one, face-to-face meeting that typically took thirty to sixty minutes. A flyer that they were invited to take home with them reiterated this information. Moreover, participants also received daily text messages that reminded them of their earnings. Similarly, in Charness & Gneezy (2009), incentive information was conveyed directly to participants in an in-person meeting.

But what happens when incentives are not so conspicuously "marketed"? Arguably, the incentives that we encounter in daily life – the fitness subsidy offered by an employer, for example – tend to be buried in fine print on insurance plans or government websites. Given prior evidence of their success in influencing behavior, one might expect incentives to be sufficiently motivating to overcome the attention problem. Yet, as suggested by recently-developed economic models, people often fail to pay attention to consequential financial information, even when that information is freely available. For example, Grubb (2015) proposes inattention as an explanation for actions taken recently by the Federal Communications Commission (FCC) to reduce "bill shock," or unexpected overage fees in the cell phone market. Even though cell phone usage data are freely available to customers at any time, concerns that consumers frequently faced bill shock caused the FCC to compel U.S. carriers in 2011 to begin sending usage alerts to customers when they had exceeded data limits and were thus at risk of experiencing bill shock (FCC, 2015). Similarly, Hanna, Mullainathan, and Schwartzstein (2014) show that people sometimes fail to learn not due to lack of data, but due to a failure to notice the data in the first place.

Consistent with the idea that people may be inattentive to information about incentives, past research has provided initial evidence that information conspicuousness may influence behavior. For example, salience influences individuals' responses to taxation (Chetty et al., 2009; Finkelstein, 2009; Taubinsky & Rees-Jones, 2017), purchasing decisions (Bordalo, Gennaioli, &

Shleifer, 2013; Busse, Lacetera, Pope, Silva-Risso, & Sydnor, 2013), water usage (Tiefenbeck et al., 2018), and reactions to college rankings (Luca & Smith, 2013). Conversely, consumers' judgments and decisions can be affected by irrelevant but salient information (Dohmen, Falk, Huffman, & Sunde, 2006). Although these findings relate to general consumer decision-making, no work to date has considered how making financial incentives salient affects people's propensity to engage in the target behavior. Therefore, we test the hypothesis that the effectiveness of incentives depends not simply on incentives' presence or size, but sometimes more crucially on whether they are presented conspicuously. In doing so, we offer the first experimental evidence that incentive conspicuousness – over and above awareness that incentives exist – affects the capacity for incentives to shift behavior.

In addition, in this paper, we explore the nature of the inattentiveness driving people's apparent unresponsiveness to inconspicuous incentives. We test whether inattentiveness to consequential incentives may be driven at least in part by a kind of apathy – a failure to seek information on how to improve one's outcomes, *despite* the knowledge that such information exists and *even when* such information can be obtained at the mere click of button. Extensive behavioral science research dating back to the mid-1900s (Simon, 1947) suggests that people often satisfice rather than optimize; as a result, people commonly choose default options, because they offer the path of least resistance (Samuelson & Zeckhauser, 1988). We hypothesize that this same kind of apathy results in a resistance to exert effort to uncover incentive information, rendering inconspicuous incentives ineffective. We offer a strong test of this hypothesis, by examining whether it is supported when the amount of effort needed to uncover incentive information is minimal.

Finally, we explore whether or not people accurately anticipate others' failure to seek such consequential incentive information. Recent research suggests that people are not only often naïve about their own biases (O'Donoghue & Rabin, 2011; Rogers & Milkman, 2016), but also those of others. For instance, resarch on "default neglect" has demonstrated that policymakers fail to appreciate how important it is to set up defaults to favor whatever option they hope others will select because they do not anticipate status quo bias (Zlatev, Daniels, Kim, & Neale, 2017). Further, people tend to over-estimate the degree to which others will respond to extrinsic incentives and under-estimate the degree to which they will respond to intrinsic incentives – an inaccurate belief that can lead policymakers to design sub-optimal incentive schemes (Heath, 1999). We hypothesize that people may similarly under-appreciate the importance of making incentives conspicuous because of inaccurate lay theories about human rationality (Heath, 1999; Van den Steen, 2004; Zlatev et al., 2017), which may explain why it is easy to identify instances in which policymakers have failed to invest in making incentives salient.

Taken together, our results suggest that it is not enough to simply inform people of the possibility that useful information about incentives exists; rather, one must make such information conspicuous. But our results also point to a barrier to doing so: policymakers may underestimate the importance of making incentive programs conspicuous due to inaccurate lay theories of how others will respond to inconspicuous incentives. Participants in our experiments incorrectly assume that others will proactively seek information on incentive programs. We discuss the implications of these findings for the numerous incentive programs designed by managers and policymakers with the objective of producing behavior change, many of which offer inconspicuous incentives and are thus likely less effectual than they could be.

The present investigation, in which we manipulate incentive conspicuousness, is related yet distinct from prior work showing that low-touch reminders can increase follow-through on desirable behaviors that are often forgotten (Ericson, 2017; Karlan, McConnell, Mullainathan, & Zinman, 2016; Karlan, Ratan, & Zinman, 2014; Rogers & Milkman, 2016; Shea, DuMouchel, & Bahamonde, 1996), such as getting a vaccination or taking medications (Briss et al., 2000; Dai et al., 2017; Szilagyi et al., 2000). Although reminders may be very effective for these types of one-off behaviors – behaviors that one can easily forget to do – they are arguably less applicable to the behavior of focus in our field experiment: walking. Walking is a behavior that our study participants engage in every waking hour and are actively monitoring with wearable devices. Therefore, forgetting to walk is not a challenge for these individuals in the same way that forgetting to get a vaccination or take a medication might be. However, conspicuous incentives ensure that an opportunity to earn rewards will not go unnoticed. In this sense, we speculate that conspicuous incentives operate by drawing attention to desirable behaviors rather than targeting forgetting, as reminders do.

We propose and test the following three hypotheses in this paper:

H1: Holding awareness of incentives constant, people will be less likely to engage in an incentivized behavior when incentives are inconspicuous than when they are conspicuous.

H2: Many people are unwilling to seek information about incentives that would make them better off.

H3: People do not anticipate H2; they overestimate potential incentive beneficiaries' propensity to seek incentive information and under-invest in making incentives conspicuous.

Experiment 1: Field Evidence

Our field experiment examines the impact of a conspicuous versus inconspicuous increase in incentives on physical activity.

Methods

We randomly assigned 2,055 pedometer users who were already enrolled in a program offering small incentives (approximately $0.05 per 10,000 steps) for walking to one of two experimental conditions. All users received increased incentives (40x what they were used to receiving, or approximately $2 per 10,000 steps) for walking for two weeks and had access to information about these incentives, which they could easily track through a rewards application available on their smart phones and an online portal. All users also received weekly emails updating them on their total earnings. Because we sought to test the conspicuousness of the increased incentives (as opposed to merely making some study participants aware of the increase and leaving others in the dark), as described below, users selected for inclusion in the study were required to have a history of regularly checking their weekly earnings balance emails. Users in the *conspicuous incentives condition* received eight extra emails containing information about the new incentives on offer. Users in the *inconspicuous incentives condition* did not receive these extra emails.

Participants

Our study population was composed of users of an online platform called Achievemint, a reward platform for healthy activities powered by Evidation Health. Achievemint users can link their Fitbit pedometers with the Achievemint platform by authorizing their recorded step counts to be automatically transferred to their Achievemint account. Every time an Achievemint user takes 200 steps, he or she earns one point from the platform. Points are redeemable for cash rewards: after a user has taken 200,000 steps, he or she earns $1.00. Users receive a check for every $25 earned. Achievemint sends all users (including all study participants throughout the course of the study) a weekly update email that contains information on a user's current number of unredeemed points. As described below, our study entailed boosting these incentives by a factor of 40, and manipulating the conspicuousness of this increase.

Of the Achievemint users who had linked Fitbit devices, we selected 2,055 for our study based on three eligibility criteria. First, users who had been part of a previous study of incentives for exercise were excluded. Second, users whose historical usage data indicated that they opened fewer than one email per month from Achievemint were excluded. This helped to ensure that users in the study actively checked their emails from Achievemint (in particular, ensuring that even those in the inconspicuous incentives condition would have easy access to information about the new incentives they were earning because they regularly opened their earnings balance emails from Achievemint). Third, users whose historical usage data indicated that they were above the 70th percentile for mean daily steps were excluded because these very active individuals did not need to increase their daily step counts.

Based on our sample size and these exclusion criteria, we had 90% statistical power to detect a 15% difference between conditions in step counts over a two week period using two-sided t-tests. Our experimental protocol was approved by the Institutional Review Board of a U.S. university. A waiver of informed consent was approved per Federal regulations (45 CFR 46.116(d)) given that the study was minimum risk, did not adversely affect the rights and welfare of participants, and could not be practicably carried out without the waiver.

Procedures

Achievemint users meeting eligibility requirements were randomly assigned to one of two experimental conditions: a conspicuous incentives condition or an inconspicuous incentives condition. All study participants received the same incentive to walk: the number of Achievemint points that they earned per step was multiplied by 40 for two weeks (that is, for every 200,000 steps, they earned $40 instead of the usual $1). Throughout the course of the study, participants in the conspicuous incentives condition ($n = 1,027$) received a low-touch email "campaign" consisting of a kickoff email (described below) plus an email every other day detailing the duration and magnitude of these increased incentives. These emails were in addition to Achievemint's standard weekly update emails containing the given user's point balance, which all participants received. Participants in the inconspicuous incentive condition ($n = 1,028$) were made aware of the increased incentives through their weekly points balance emails[2], in addition to having anytime access to this information via the online portal. Thus, in both experimental

---

[2] Unfortunately we were unable to obtain participants' email open rates during the intervention from our field partner. However, we do have data on participants' pre-intervention propensity to open these weekly emails. This propensity did not differ between conditions, suggesting that participants in the inconspicuous condition were just as likely to open these weekly emails as those in the conspicuous condition. See Table 1.

conditions participants had regular, on demand, easy access to an up-to-date report on the increased incentives that they were earning. Critically however, participants in the inconspicuous incentives condition did not receive the low-touch campaign designed to increase the conspicuousness of these increased incentives.

Participants received increased incentives for walking for two weeks from January 27, 2015 through February 9, 2015. Participants in the conspicuous incentives condition received a kickoff email on January 26, 2015, the day before the start of the increased incentives. This email featured the subject line: "New Program to Encourage You to Walk (earn Bonus Points)." The contents of this email (depicted in Appendix, Figure A1) explained to participants that they had been enrolled in a program to increase their walking. It showed a calendar with point multipliers highlighted on each day when they would earn increased incentives (every day for the next two weeks). Conspicuous incentive condition participants also received email notifications about the program every other day for its duration (seven additional emails on days 1, 3, 5, 7, 9, 11, and 13 of the experiment), which contained all of the same information including the schedule of incentives depicted on a calendar. A notification email is depicted in Appendix, Figure A2.

No participants opted out of the incentive program. However, as we discuss further in the results section, some participants did not record daily steps on some days, either because they failed to wear Fitbits, or because they did not sync their Fitbit data with Achievemint. Other participants may have worn Fitbits only for a brief portion of a given day.

Statistical Analyses

Our outcome variable of interest was daily steps taken. Participants' daily steps were tracked for three weeks before the intervention, two weeks during the intervention, and three weeks after the intervention. We tested for evidence of differential responses to the increased incentives offered by comparing mean daily steps between conditions during and after the end of the intervention period.

Our statistical analysis strategy was a difference-in-differences approach. Like other standard difference-in-differences analyses (e.g., Pope & Pope, 2015), our analyses include covariates for experimental condition, temporal indicators (during-, and post-intervention), and the interaction of experimental condition and temporal indicators. We chose difference-in-differences analysis rather than simple comparison of groups because it is effective at comparing the time changes in the means between groups, accounting for both group-specific and time-specific effects (Wooldridge, 2010). This analytical strategy therefore enables comparisons between experimental groups (conspicuous and inconspicuous) at different times (during and after the intervention) within person, providing considerably more statistical power than a mere comparison of means. The coefficients of interest in regressions of this form are the interactions of experimental condition and temporal indicators. These coefficients measure the effect of the treatment at a given time period (Wooldridge, 2010). Importantly, our study uses random assignment to experimental condition, so experimental groups are the same in expectation (Athey & Imbens, 2006). Further, there were no significant differences in the observable pre-treatment characteristics of the two experimental groups (see Table 1).

Our analyses rely on ordinary least squares regressions with person-day-of-the-week fixed effects and date fixed effects. To account for different walking levels on different days (due to weather, day of week, and seasonality), we include fixed effects for each date observed in the

dataset, instead of time period dummies. It would be natural to include user fixed effects in an analysis such as this one to account for different individual activity levels. However, in addition to varying in overall activity levels, individuals also vary substantially in their activity patterns throughout a given week. For example, some individuals may take a recreational Sunday walk, while others walk more during the week as part of a commute. Individuals work on different days and have different exercise and gym attendance habits. To account for these idiosyncrasies, we employ fixed effects that are more specific than user fixed effects. Specifically, we include fixed effects for an interaction of user with day of the week (Monday, Tuesday, etc.): seven fixed effects per individual. We also cluster standard errors at the user level, to account for possible serial correlation within a user's daily steps over time. This analysis strategy increases our ability to identify experimental treatment effects.

The following is the ordinary least squares (OLS) regression equation used to estimate the coefficients shown in Table 2 as part of our difference-in-differences strategy.

(1)	$daily\_steps_{it} = \beta_0 + \alpha_1 user\_x\_day\_of\_week_{it} + \alpha_2 day_t$ $+ \beta_1 conspicuous\_incentives_i \times during\_treatment_t$ $+ \beta_2 conspicuous\_incentives_i \times 0to2weeks\_post\_treatment_t$ $+ \beta_3 conspicuous\_incentives_i \times more\_than\_2weeks\_post\_treatment_t +$ $\mathcal{E}_{it}$

In this equation, *conspicuous_incentives*$_i$ is an indicator variable taking on a value of one if an individual, *i*, was in the conspicuous condition and zero otherwise; *day*$_t$ is a fixed effect for each day included in the data; *user_x_day_of_week*$_{it}$ is a fixed effect for user-day-of-week, and the other variables represent 0-1 indicators for whether an observation occurred during, within the first two weeks after, or during the third week after the intervention. The β_1, β_2, and β_3 coefficients therefore measure the differences between conspicuous incentives condition participants and inconspicuous incentive condition participants during, shortly after, and long after the experimental intervention.

In addition to our primary analyses, we conducted supplemental analyses to compare participants in our experimental groups with participants in a matched holdout group. We selected the matched holdout group from the same pool of Achievemint users from which the experimental participants were selected (those with linked Fitbit accounts). To select a holdout group that closely matched the experimental group in observable characteristics, including demographics, we selected Achievemint users who were excluded from our experiment only because they had been participants in a previous experiment, but who would not have been excluded otherwise (because they met all other inclusion criteria as experimental participants).[3]

The following is the OLS regression equation we used to estimate the effect of (a) an inconspicuous increase in incentives relative to receiving the standard, smaller incentives, and (b) conspicuously increased incentives relative to inconspicuously increased incentives. Coefficients from this regression analysis are shown in Table 4 as part of our secondary difference-in-differences analysis. This regression's standard errors were also clustered by user and we again include user-day-of-week fixed effects as well as day fixed effects.

(2)	$\begin{aligned} daily\_steps_{it} = {} & \beta_0 + \alpha_1 user\_x\_day\_of\_week_{it} + \alpha_2 day_t \\ & + \beta_1 in\_experiment_i \times during\_treatment_t \\ & + \beta_2 in\_experiment_i \times 0to2weeks\_post\_treatment_t \\ & + \beta_3 in\_experiment_i \times more\_than\_2weeks\_post\_treatment_t \\ & + \beta_4 conspicuous\_incentives_i \times during\_treatment_t \\ & + \beta_5 conspicuous\_incentives_i \times 0to2weeks\_post\_treatment_t \\ & + \beta_6 conspicuous\_incentives_i \times more\_than\_2weeks\_post\_treatment_t \\ & + \mathcal{E}_{it} \end{aligned}$

In this equation, $in\_experiment_i$ is an indicator variable taking on a value of one if an individual, i, was in the experiment and zero otherwise, $conspicuous\_incentives_i$ is an indicator variable taking on a value of one if an individual, i, was in the conspicuous condition and zero otherwise,

---

[3] Because difference-in-differences analyses can encounter problems when comparing groups with underlying differences, it is possible that our analyses using Equation 2 (found in Table 4) are biased. However, our matched holdout group was selected specifically to be similar to the experimental group in observable characteristics.

day$_t$ is a fixed effect for each day included in the data, *userxday_of_week$_{it}$* is a fixed effect for

user-day-of-week, and the other variables represent 0-1 indicators for whether an observation

occurred during, within the first two weeks after, or during the third week after the intervention.

Like Equation 1, Equation 2 represents a difference-in-differences analysis strategy. In our

analyses, we also directly compare the effect of the conspicuous incentive condition to the

holdout group using Wald tests of the coefficients from Equation 2 (β_4- β_1, β_5- β_2, and β_6- β_3).

Results

Primary results

Figure 1 shows the differences in mean daily steps taken by participants in the

conspicuous and inconspicuous incentives conditions during and after our intervention. As

predicted, Figure 1 shows that users in the conspicuous incentives condition had a higher mean

daily step-count both during and after the intervention relative to the inconspicuous incentives

group. To test the significance of these differences, we turn to regression analyses. We are able

to identify the effects of our intervention with considerable precision by controlling for

individual differences in pre-intervention propensity to walk while identifying differences

between our experimental groups in during- and post-intervention behavior. As described

previously, our analysis uses a difference-in-differences regression strategy, which in general

enables comparisons between groups and over time. For our research, the difference-in-

differences strategy enables us to compare the mean daily steps of the participants in the

conspicuous incentives group with the mean daily steps of participants in the inconspicuous

incentives group during, shortly after, and long after our experimental intervention.

Table 2 shows the results of OLS regressions predicting daily steps with fixed effects for each participant on each day of the week as well as date fixed effects. The primary predictor variables of interest are interactions between an indicator for our conspicuous incentives condition and indicators for each of the time periods studied during and post-intervention (two weeks during the intervention, two-weeks immediately post-intervention, and three weeks post-intervention). Standard errors are clustered by participant and reported in parentheses. The primary coefficients of interest in these regressions estimate the difference between conspicuous and inconspicuous incentives groups' mean daily steps during each time period of interest during and post-intervention after absorbing: (a) average individual differences in exercise as a function of day of week, (b) average pre-treatment differences in participants across condition (since the same participant is always in the same experimental condition), and (c) average differences in participant exercise by date.

Table 2, Model 1 reports the results of this regression estimated on an intent-to-treat basis. As predicted, conspicuous incentive condition participants took significantly more daily steps during the intervention than inconspicuous incentive participants (an estimated 367 extra daily steps, representing a 7% increase, $p < .001$). Further, our regression estimates indicate that participants in the conspicuous incentives condition took more daily steps than those in the inconspicuous condition for the two weeks after our intervention (an estimated 332 extra daily steps, representing a 6% increase, $p < .01$), although the regression estimated effect dissipated three weeks after our intervention (an estimated 146 extra daily steps, *NS*).

Addressing alternative interpretations

One concern about the above results is that they could be driven by differential use of Fitbits rather than differential exercise if conspicuous incentives prompted participants to wear

their Fitbits more often than participants in our inconspicuous incentives condition without actually changing their walking habits. One way of addressing this concern is to examine differences in attrition by condition. One measure of attrition in Fitbit usage could be indicated by a user recording zero steps for a given day. Attrition could also be indicated by a user recording a small, nonzero number of steps, for example if he or she only wore a Fitbit for a brief part of a day. A previous study of walking with an adult participant pool consisting of several thousand people with no attrition found that participants never took fewer than 2,000 steps per day (Hirvensalo et al., 2011). We therefore used a very conservative 2,000 steps as a cutoff for defining a participant an "attriter": days on which a given user recorded fewer than 2,000 steps were regarded as a failure to properly wear and sync Fitbits (note that all reported results are stronger if we instead define attrition as 0 steps per day). Table 1 shows the numbers of attriters (fraction of people who recorded < 2000 steps on any day) in each experimental condition during the 21 days pre-intervention. We performed two-sample proportion tests to determine whether the conspicuous and inconspicuous incentives groups had different attrition rates before our intervention. We see (in Table 1) no significant differences at the .05 significance level between the attrition rates in the experimental conditions before our experimental intervention.

However, to ensure the robustness of our regression results to attrition, we replace our intent-to-treat analyses with two additional, more conservative tests of our hypothesis in Table 2. In one analysis (Model 2), we delete all person-day observations from both experimental conditions that show fewer than 2,000 daily steps logged. In another analysis (Model 3), instead of deleting person-day observations with fewer than 2,000 daily steps, we replace such observations with the average of the relevant individual's pre-intervention daily step counts stratified by day of week and excluding days with fewer than 2,000 steps.

The coefficient estimates reported in Table 2 in Models 2 and 3 provide more evidence that conspicuous incentives boost step counts significantly and that the observed effects reported in Model 1 are not artifacts of differential attrition. According to the estimates from Model 2, which again control for participant-day-of-the-week fixed effects and date fixed effects, non-attriting participants in the conspicuous incentives condition took more steps than participants in the inconspicuous incentives condition during the intervention (an estimated 229 extra daily steps, representing a 4% increase, $p < .01$) and up to two weeks after the intervention (an estimated 211 extra daily steps, representing a 4% increase, $p < .05$), but not three weeks after the intervention (an estimated 96 extra daily steps, representing a 2% increase, *NS*). The results of Model 3 are similar to those in Model 2: participants in the conspicuous incentives condition took more steps than participants in the inconspicuous incentives condition during the intervention (an estimated 167 extra daily steps, representing a 3% increase, $p < .01$) and up to two weeks after the intervention (an estimated 177 extra daily steps, representing a 3% increase, $p < .05$), but not three weeks after the intervention (an estimated 65 extra daily steps, representing a 1% increase, *NS*). Although the coefficient estimates in these models are both smaller than those estimated in Model 1, the effects of the intervention during the intervention period and for the two weeks post-intervention remain large, positive and significant, indicating that differences between experimental groups cannot be explained by differential attrition.

Differential awareness of the incentives is another potential alternative interpretation of our results. It is possible, for example, that the treatment not only increased the conspicuousness of the incentives, but also awareness of their existence. Indeed only participants in the conspicuous condition received the kick-off email, which introduced the increased incentives. Although all participants could access this information by logging into their platform, because

those in the conspicuous condition also received a kick-off email describing the program, it is possible that a greater proportion of participants in this condition were aware of the incentives, and thus, that differences in awareness of the increased incentives that is responsible for the observed effect.

One way to assess whether differences in incentive awareness drove our results is to evaluate whether they remain consistent among a subset of participants likely to be aware of the incentive program. For those in the inconspicuous condition, the primary means of accessing the incentive information was via opening weekly emails. Although we were unable to obtain individual-level email open rates from the intervention period from our field partner, we assessed the effect of the conspicuousness intervention for participants with varying levels of historical email open rates. Those with high (pre-intervention period) email open rates were probably particularly likely to have opened the weekly summary emails during the intervention. Participants with high email open rates in the conspicuous and inconspicuous conditions alike were particularly likely to have been aware of the increased incentives. Table 3 shows the results from the same OLS regressions predicting daily steps shown previously in Table 2 Model 1, but this time stratified by email open rate quartile. As shown in Column 4, the effect of increased incentive conspicuousness was a significant predictor of daily steps even among those most likely to have been aware of the incentive information - the top quartile of email openers.

Comparing the effect of increasing incentives to that of making this increase conspicuous

All results reported thus far have compared users across our two experimental conditions (conspicuous incentives and inconspicuous incentives). Both experimental groups however, received the same increased incentives during our two week intervention period. Thus, while we

have measured the effect of making this increase conspicuous, we have not measured the effect of the increased incentives themselves.

Using a difference-in-difference estimation strategy, we now turn to comparing the change in pre- versus during- and post-intervention behavior of participants in our inconspicuous incentives group with the behavior of participants in a holdout group who were excluded from our experiment only because they had previously participated in another, unrelated experiment, but who otherwise would have been eligible for participation. Because this group met the study's eligibility criteria based on observable characteristics, it provides an ideal control group for comparisons with experimental participants in a difference-in-differences analysis.

Table 4 shows the results of ordinary least squares regressions predicting daily steps with fixed effects for user-day-of-week and the calendar date. The key predictor variables are interactions between experimental status (in experiment, in conspicuous incentives condition) and different time periods (during- and post-intervention). Standard errors are clustered by participant and reported in parentheses. As in Table 2, the results from three models are included: Model 1 includes all users, whereas Model 2 deletes all person-day observations with fewer than 2,000 daily steps logged and Model 3 replaces such observations with the average of the relevant individual's pre-intervention daily step counts stratified by day of week (and excludes pre-intervention days with fewer than 2,000 steps).

The first three coefficients in these regressions estimate the difference between inconspicuous incentive participants' and holdout users' mean daily steps at each time period—they estimate the effects of the increased incentives after controlling for user-day-of-week and calendar date fixed effects. The second three coefficients estimate the difference between conspicuous incentive participants' and inconspicuous incentive participants' mean daily steps at

each time period—they estimate the marginal effect of conspicuousness on responses to these increased incentives, again after controlling for user-day-of-week and calendar date fixed effects.

The results in Table 4 indicate that inconspicuous incentive group participants' behavior did not differ from the behavior of holdout group participants during the intervention, though they may have differed after the intervention in the opposite of the intended direction. Specifically, at 0–2 weeks post intervention, one of the models shows a statistically significant difference such that inconspicuous incentive group members exercised less than members of the holdout group. Given that this result was not predicted a priori, and that it only surfaced in one of the models, we speculate that it is an anomaly; however, future research could ascertain whether this pattern replicates (and if so, why). More importantly, all three models in Table 4 converge on the conclusion that although making the increase in incentives conspicuous was effective in changing behavior, increasing the incentives did not, in and of itself, meaningfully alter exercise.

We further assessed the effect of conspicuously increased incentives by directly comparing participants' activity in the conspicuous condition to the holdout group via Wald tests, also reported in Table 4. As predicted, participants in the conspicuous incentive condition took significantly more daily steps during the intervention than those in the holdout group, although this effect did not persist after the intervention.

Robustness tests

Although random assignment to conditions was successfully balanced on static traits, as reported in Table 1, we also conducted parallel trends analyses to test for pre-treatment equivalence in the *trajectory* of daily step counts between groups. The results are summarized in Supporting Information Table S1. The assumption of parallel pre-treatment trends is not well-supported in SI Table S1, suggesting that random assignment was unsuccessful on this

observable dimension of our sample. A large source of noise in our dataset is attrition, or days

when participants failed to wear their pedometers and record steps, and we suspected this noise

could be driving the imbalanced parallel trends detected pre-treatment. In fact, a comparison of

the proportion of pre-intervention days when participants took fewer than 2,000 steps revealed a

significant difference between conditions ($M_{conspicuous} = 0.25$, $SD_{conspicuous} = 0.01$, $M_{inconspicuous} =$

0.22, $SD_{inconspicuous} = 0.01$; $t(2053) = -2.74$, $p < .01$).

To test the possibility that participants who wore their pedometers unreliably introduced

noise that led our data to fail the parallel trends test, we re-ran our parallel trends analysis

excluding any participants whose steps were missing any day pre-treatment, and when we do

this, our data indeed supports the assumption of parallel pre-treatment trends. SI Table S2 shows

that the parallel trends test is well-supported when we focus on this sub-population. SI Figure S1

depicts pre-treatment trends for all of our data and for the sub-sample including only users who

wore pedometers every day pre-treatment. This figure visually depicts that parallel trends were

present in the sub-population of participants who wore pedometers each day pre-treatment but

not in the full population. To ensure imbalances in parallel trends could not be responsible for

our primary findings, we repeated all of our analyses only including the subset of participants

who had no missing step count data pre-treatment (and whose parallel trends were balanced, as

shown in SI Table S2). Our results are presented in SI Table S3 and SI Table S4, and are robust

in this sub-population, suggesting that imbalanced step count trajectories cannot account for our

findings.

In addition to the analyses reported above, we conducted a variety of robustness checks

on our results (see Supporting Information). Specifically, we performed the following robustness

checks: (1) we clustered data by user-day-of-week, using data only on experimental participants

(SI Table S5); (2) we clustered data by user-day-of week, using data on experimental participants together with the matched holdout group (SI Table S6); (3) we clustered data by user, winsorizing step counts at the 99% level, using data only on experimental participants (SI Table S7); (4) we clustered data by user-day-of-week, winsorizing step counts at the 99% level, using data only on experimental participants (SI Table S8); (5) we clustered data by user, winsorizing step counts at the 99% level, using data on experimental participants together with the matched holdout group (SI Table S9); and (6) we clustered data by user-day-of-week, winsorizing step counts at the 99% level, using data on experimental participants together with the matched holdout group (SI Table S10). In each of these models, users in the conspicuous incentives condition take more daily steps than users in the inconspicuous incentives condition both during and up to two weeks after the experimental intervention, with $p < .05$ for both of these coefficients in each model.

Discussion

This field experiment is consistent with the premise that incentives alone can be insufficient to change behavior: physical activity levels were no higher among participants who were inconspicuously offered a 40x increase in incentives than among those in a holdout comparison group who received the standard, much smaller, incentives. By contrast, a simple and inexpensive email campaign designed to make these increased incentives more conspicuous was sufficient to unlock their power to meaningfully change behavior for the better. Thus, Experiment 1 suggests that an increase in incentives will impact behavior only when that increase is conspicuous. In our next experiment (Experiment 2), there was no pre-intervention

period in which participants were offered a smaller incentive. Thus, in Experiment 2, we test whether making incentive information conspicuous – as opposed to not conspicuous – affects behavior.

Experiment 2: Controlling Information Exposure

Building on our field study, in an online experiment we test a more conservative manipulation of incentive conspicuousness. Specifically, all participants were informed about incentives intended to influence their behavior. However, for half of participants, we made this information more conspicuous by presenting it in a more prominent font.

This design addresses two limitations of our field study. First, it addresses the issue of differential awareness of incentives because incentive information was presented to everyone (whereas in our field study, participants had to check their earnings online or read their weekly emails about incentives to notice a change in their earnings). Second, it addresses the possibility that our manipulation changed participants' beliefs about the importance of engaging in the incentivized behavior because everyone was exposed to the same information about incentives.

Experiment 2 was a 2 condition between-subjects design. As specified in our pre-registration (https://aspredicted.org/blind.php?x=63zy8z), we predicted that participants would be more likely to engage in an incentivized behavior when assigned to an experimental condition in which incentives were made more conspicuous.

Method

Three hundred and five participants were recruited online from Prolific Academic (47% male, M_{age} = 30.3 years, SD = 10.3) and were told upfront that they would be given a $0.60 payment for completing the survey. They were also told that they would have the opportunity to earn an additional bonus payment of up to $0.16. All participants were presented with the following brief description of the study procedures, which clearly conveyed information on the (somewhat peculiar) way they could earn this additional $0.16 bonus payment:

> During this study, you will answer several questions about yourself and your opinions. It is important that you share your true opinions as there are no right or wrong answers.

> In part I of this study, you will share your opinion about a picture.

> In part II of this study, you will answer some survey questions. During Part II of this study, you will have the opportunity to earn an additional bonus. Four of the multiple choice questions will include an option to select 'other' and write in an open-ended response. You can earn a four cent bonus for each of these questions by selecting the 'other' multiple choice answer and writing in the phrase "seagulls fly over the sea." If you wish to write in your own response to this question, you may write it after this bonus phrase.

This information was presented on a single page with no other information – i.e., the information about how to earn the bonus incentive was clearly conveyed to all participants. For half of participants, we made this information more conspicuous; for these individuals, the sentence "You can earn a four cent bonus for each of these questions by selecting the 'other' multiple choice answer and writing in the phrase "seagulls fly over the sea." was presented in a

more conspicuous way (i.e., font was larger, bold, and highlighted; more white space between text; see Appendix as Figure A3 for a screenshot).

Next, all participants responded to a 50-item multiple choice opinion survey about politics and news, with items sourced from the Pew Research Center, a nonpartisan research organization (www.pewreserach.org) (e.g. "Thinking just about science news, how often do you read, watch or listen to news about science?" (1 = *nearly every day*, 2 = *a few times a week*, 3 = *a few times a month*, 4 = *less often*). Four items within this survey included an "other" multiple choice option with an open-ended text box next to it, where participants could write the bonus phrase. The survey concluded with questions about participants' age and gender.

Results and Discussion

Although all participants viewed the information about the bonus incentive, those in the more-conspicuous condition were significantly more likely to engage in the incentivized behavior (i.e., to write in the bonus phrase) than those in the less-conspicuous incentive condition ($M_{more\text{-}conspicuous}$ = 1.89 bonus phrase write-ins, of four opportunities, $SD_{more\text{-}conspicuous}$ = 1.77; $M_{less\text{-}conspicuous}$ = 1.27 bonus phrase write-ins, $SD_{less\text{-}conspicuous}$ = 1.69), $t(303) = 3.09$, $p = .002$, $d = .35$. Participants in the more-conspicuous condition earned an average of \$0.025 more in bonus pay – a 48% increase relative to the average bonus pay earned in the less-conspicuous condition.[4]

---

[4] Six participants wrote in variations of the requested bonus phrase (e.g. "seagulls fly over the ocean," "the seagulls fly over the sea"); these participants were considered as having engaged in the incentivized behavior (our results are similar when these entries are counted as incorrect).

These results support our prediction that conspicuous incentives increase the uptake of an incentivized behavior in a new setting, highlighting the robustness of our field study results. They also address potential concerns that the results of the field experiment are specific to the behavior in question (walking), or to the experimental context, or the interaction between them. Although Experiments 1 and 2 point to the importance of making incentives conspicuous, this is not to say that conspicuous incentives are always effective; for example, their impact may decrease over time, as people adapt to them.

Experiment 3: Valuing Conspicuous Incentive Information

This experiment had two goals. First, we sought to assess people's demand for information about incentives. Second, we explored whether people anticipate others' propensity to seek (or not seek) information about incentives.

The experiment was a three condition between-subjects design: in two of the three conditions, similar to Experiment 2, participants could earn a bonus incentive for engaging in a certain target behavior. Between-subjects, we randomized these participants to either be exposed to the incentive information by default (*default exposure condition*), or to have the choice of whether to seek this information (in which case we instructed them to click on a link to view this information, *opt-in exposure condition*). Participants in a third, *policymaker condition* were asked – in an incentive compatible manner – to predict whether participants in the *opt-in exposure* condition would seek the incentive information.

As indicated in our preregistration https://aspredicted.org/blind.php?x=7t3xs5, we predicted that participants in the default exposure condition would be more likely to engage in the target behavior relative to those in the opt-in condition. Second, we predicted that policymakers would act as if they believe that prospective beneficiaries can be left to their own devices to seek valuable incentive information.

Method

Three hundred and one participants were recruited online from Amazon Mechanical Turk. Per our pre-registration, we excluded 24 participants (default exposure: n = 3; opt-in exposure: n = 11; policymaker: n = 10) because they indicated that they had had technical difficulties with the study, resulting in a final sample of 277 participants (54% male, M_{age} = 36.8 years, SD = 11.4). Participants were guaranteed a $0.50 payment and told they would have the opportunity to earn a bonus payment of up to $0.14.

Similar to Experiment 2, all participants were presented with an introductory page on which they were told that during the study they would be asked to answer questions about themselves and their opinions. The rest of this introductory page differed by experimental condition. Participants in the default exposure condition were told:

> *During this study, you will have the opportunity to earn an additional bonus. 3 of the questions will include an option for you to fill in an open-ended text box. For each of these questions, you can earn an additional $0.04 each time you write the words "seagulls fly over the sea" in the open ended response.*

Then, they viewed an example of an open-ended response question with the bonus phrase written in.

Participants in the opt-in exposure condition were told:

During this study, you will have the opportunity to earn an additional bonus. Instructions for how to do so are at this <u>*link*</u>*.*

Upon clicking the link, the same text that was presented automatically in the default exposure condition appeared.

Participants in the policymaker condition were told:

You will have the opportunity to earn up to a $0.12 bonus based on the actions of another mTurker completing this survey, with whom you will be randomly paired. 3 of the questions in this survey will include an option to fill in an open-ended text box. For each of these questions, you can earn an additional $0.04 each time THE MTURKER YOU ARE PAIRED WITH writes the words "seagulls fly over the sea" in the open ended response.

Participants in the policymaker condition were also told that the person with whom they would be paired would earn an equivalent bonus for themselves.

Next, participants in the policymaker condition were told that the mTurk worker with whom they would be paired would be drawn from the opt-in exposure condition (which was termed "Group A"), and were shown a screenshot of the introductory page from the opt-in exposure condition. Then, policymakers were told:

Instead of being linked with a participant from Group A, you may elect to pay some of your compensation in order to be paired with a randomly selected mTurker from Group B. MTurkers in Group B will have the instructions displayed to them automatically - i.e., they won't have to click on the link to see the instructions; the instructions will by default be displayed to them, without having to click on a link.

Policymakers were then required to correctly answer two comprehension questions to ensure they understood these instructions; these questions assessed knowledge that their bonus pay would be based on the performance of a participant from the opt-in exposure condition, unless they bid to be paired with someone from the default exposure condition.

Next, policymakers completed a Becker-DeGroot-Marshak procedure (Becker, DeGroot, & Marschak, 1964) to assess their willingness to pay to be paired with an mTurker from the default exposure condition (full stimuli posted at https://osf.io/tnxvp/?view_only=c1bda95170704fe3986046927c05cd40).

Next, all participants responded to the same 50-item multiple choice opinion survey as in Experiment 2. Three items included an "other" multiple choice option wherein participants could write the bonus phrase. The survey concluded with demographic questions and a comprehension check in which participants were tasked with correctly identifying, amid a set of distractor statements, that they could earn a bonus based on whether "seagulls fly over the sea" was entered into the "other" response option. They were told (truthfully) that they would earn an additional $0.02 for answering this question correctly. We expected participants in the opt-in exposure condition to be significantly less likely to pass this check, as we anticipated that many would not seek the incentive information. Finally, participants indicated whether they had experienced any technical difficulties as well as their age and gender.

Results and Discussion

Although doing so was nearly costless, only 26% of participants (i.e., 22 out of 86) in the opt-in exposure condition sought the incentive information (i.e., clicked the link).

Correspondingly, participants in the opt-in exposure condition were less likely to engage in the incentivized behavior (i.e., write in the bonus phrase) (M_{opt-in} = 0.50 bonus phrase write-ins, of three opportunities, SD_{opt-in} = 1.06) relative to those in the default exposure condition ($M_{default\ exposure}$ = 1.56 bonus phrase write-ins, $SD_{default\ exposure}$ = 1.28), $t(182)$ = 6.08, $p < .001$, d = 0.90. The difference translates to a $0.042 reduction in bonus pay: on average, participants in the opt-in exposure condition earned 68% less incentive pay relative to participants in the default exposure condition.[5]

Next, we address whether policymaker participants appreciated the value of mandating exposure to incentive information. Only 14% of policymakers (i.e., 13 out of 93) chose to be paired with someone from the default exposure condition. Put differently, 86% of our policymakers were unwilling to spend so much as a penny to ensure that their partner was exposed to the incentive information.[6] On average, policymakers were willing to pay a mere $0.005 to guarantee their partner's exposure to the incentive information – a sum significantly lower than the $0.042 value of mandated exposure ($t(92)$ = 25.38, $p < .001$, d = 2.64). Looking only at the 13 policymakers who were willing to pay for mandated exposure, they were willing to pay $0.036 for a partner from the default exposure condition ($t(12)$ = 1.25, p = .236, d = .35).

Thirty percent of participants in the opt-in exposure condition correctly answered the comprehension check question, compared to 87% of those in the default exposure condition

---

[5] Twelve participants wrote in variations of the requested bonus phrase (e.g. "seagulls fly over the ocean," "the seagulls fly over the sea"); these participants were considered as having engaged in the incentivized behavior (our results are similar when these entries are counted as incorrect).

[6] In a different preregistered version of the policymaker condition (https://aspredicted.org/blind.php?x=ir2jn9), we simply asked policymakers whether they would like to have their partner automatically shown the incentive information (rather than click on a link to access it), without introducing any fee for this option. If policymakers appropriately value incentive conspicuousness, all participants would choose to have a partner in the default-exposure condition when doing so was free. Yet, even when the option was costless, only 88% of policymakers chose to have their partner view the incentive information automatically (a binomial test indicated this percentage was significantly below 100%, $p < .001$).

(χ^2=61.1, $p < .001$) and 78% in the policymaker condition (χ^2=42.1, $p < .001$). Pass rates in the latter two conditions were statistically equivalent (χ^2=2.3, $p = .132$).

In sum, consistent with H2, only about a quarter of people were willing to seek information on incentives, even though this helpful information could be obtained with the click of a button. Yet, consistent with H3, policymakers do not seem to anticipate this disinterest in valuable incentive information.

General Discussion

Managers and policymakers inside and outside of government rely on incentive programs to encourage valued behaviors. While often successful, there are also notable and surprising examples of their ineffectiveness. Why don't incentives put in place to change highly elastic decisions always produce the desired response? In this paper, we provide evidence for one explanation: policymakers often fail to conspicuously promote such incentives due to inaccurate lay theories about the importance of incentive conspicuousness.

First, in a large-scale field experiment, we showed that even when increased incentives are transparently provided and easily trackable in real time, failing to make them conspicuous renders them ineffectual at shifting behavior. We extend these results with a second, lab experiment, in which we replicate these findings in another setting. Specifically, we show that even when everyone is made aware of incentives, if incentives are simply described in a more conspicuous way (i.e., in a bolder font), it increases people's propensity to engage in incentivized behaviors. Moreover, in a second lab experiment, we show that inconspicuous incentives may be

ineffective in part because people fail to seek out information about the incentives they face *even when* it can be obtained at the mere click of button. Finally, we show that people fail to anticipate others' tendency to neglect inconspicuous incentives and to seek them out incentive information at a low rate, suggesting that purveyors of incentive programs may under-invest in drawing attention to them.

Our investigation advances research on incentives in three primary ways. First, we illuminate a new, and important boundary condition to incentive effectiveness: their conspicuousness. Although prior work has illustrated that information salience shapes consumer behavior (e.g. Finkelstein, 2009; Luca & Smith, 2013; Taubinsky & Rees-Jones, 2017), it would be reasonable to expect the opportunity to earn incentives might be immune to attention problems. However, our findings suggest that when they are inconspicuous, incentives exert little if any influence on behavior. Second, we document a general apathy among consumers about uncovering incentive information. Prior experimental research on incentives has largely tested their effects in settings where information about payments was made extremely salient, leaving no role for apathy (e.g. Charness & Gneezy, 2009; Volpp et al., 2008). This was for good reason – this past work sought to test the impact of incentives, not the means by which information about incentives was delivered. Our finding that the majority of people are unwilling to engage in any effort (i.e. simply clicking a link) to uncover incentive information suggests a need for further attention to incentive information delivery to overcome consumer apathy. Finally, our work extends beyond the consumer perspective to consider the views of policymakers, or those who invest in designing incentive programs. Beyond documenting the importance of incentive conspicuousness, we show that people do not anticipate consumers' degree of apathy about uncovering incentive information. The naivete we document suggests many incentive programs

may be underutilized in part because their designers do not appreciate the importance of making the incentives themselves more conspicuous.

One symptom of such under-investment would be a general lack of awareness of incentive programs, even among their intended beneficiaries. To that end, we polled a nationally representative sample of 106 U.S. citizens (see Supporting Information, Table S11 and Table S12 for details) and assessed their knowledge of a wide range of existing U.S. Federal incentive programs. Respondents were presented with a list of 35 real subsidy programs and tasked with identifying which were offered by the U.S. government (versus only in other countries). On average, participants correctly categorized only 63% of the incentive programs (20 were U.S. programs, 15 were not). More importantly, only 66% of those who reported being eligible for a given U.S. incentive program were actually aware of it before taking our survey, and only 55% reported actually taking advantage of those subsidies for which they were eligible. Given that our field and laboratory studies suggest that inconspicuous incentives are ineffective, these survey results suggest that the U.S. government may be influencing far fewer behaviors through policy than it could be if many of its inconspicuous incentive programs were more aggressively promoted.

Opportunities for future research

We see several opportunities for future research. One topic of practical interest may be to compare the effect of making incentives conspicuous to other ways of shaping behavior. For example, how big is the effect of making incentives conspicuous relative to that of increasing the amount of those incentives? Relatedly, future work could directly compare the effect of

incentives to engage in some target behavior to the "marketing" of that target behavior (which, in our field experiment for example, could take the form of messaging to encourage people to walk). Similarly, given that tracking one's behavior can, in and of itself, spur behavior change (Goldstein et al., 2019), researchers could compare the effect of wearing a Fitbit (i.e., tracking), to that of incentive conspicuousness. Together with data on the costs of intervention, such research would be particularly relevant to policymakers, as it could be used to calculate the return on investment of interventions for behavior change.

A punchline of this research is that inconspicuous incentives are ineffective, and they are ineffective in part because people are unwilling to exert effort to learn about them. Future research could explore factors that explain the converse: why conspicuous incentives are effective. For example, conspicuously-presented information may be more memorable, enabling people to more easily recall incentives when they become applicable (Rogers & Milkman, 2011). Another possibility is that being exposed to conspicuous incentives over time, as was the case in our field study, could trigger loss aversion, spurring people to action (Kahneman, 2011) – such that failure to engage in the behavior is coded as a loss (more so than it might be when the incentives are inconspicuous). Future research could more directly consider the paths through which conspicuous incentives drive behavior, complementing our research which documents the role of apathy in explaining the failure of inconspicuous incentives.

In addition to showing that conspicuousness matters, we also found that people were generally unwilling to *seek* incentive information, and that policymakers do not seem to anticipate this unwillingness. Future research could seek to uncover the reasons underlying policymakers' apparent inability to intuit people's resistance to uncovering incentive information. One idea is that it be related to a 'curse of knowedge' type phenomenon:

policymakers, as designers of such incentive programs, are well aware of the existence of these programs. This awareness may impede policymakers' abilities to empathize with the state of would-be recipients, who may be unaware of the possible existence of such programs, and thus, fail to inquire about them. Our findings are consistent with past research suggesting that we are often unsophisticated about the degree to which others will exhibit bias (Zlatev et al., 2017; Heath, 1999). Future research exploring these ideas further would be valuable.

Finally, future research may also test the boundaries of the effects of making incentives conspicuous. For example, it could investigate whether the effects of incentive conspicuousness may diminish over time. It could be that continual exposure to conspicuous incentive information might dull its effects, or even result in recipient backlash (if, say, the incentive information distracted people from other important tasks). On the other hand, the value of repeatedly presenting a target population with conspicuous incentives could persist. Perhaps incentive conspicuousness consistently enables people to more effectively prioritize desired behaviors above other actions competing for time and attention. Future work could also consider how the time of the delivery of conspicuous incentive information influences its effectiveness. Although our field experiment did not vary the time of day in which the conspicuous incentive information was relayed, incentive information timing may influence the effect of conspicuousness on behavior. We suspect that presenting incentive information conspicuously at times in which the behavior is most likely to be undertaken (e.g. at dinner time for nutrition incentives) may amplify its effects. Finally, future research could test how the framing of the incentive information may modulate the effect of making that information conspicuous – for example, testing conveying an increase in incentives in relative versus absolute terms.

Concluding Comment

      In recent years there has been increased interest from academics and policymakers in using behavioral insights to drive behavior change (Chapman, Li, Colby, & Yoon, 2010; Loewenstein, Asch, & Volpp, 2013; Shafir, 2012). From the 2015 Obama directive to use such insights in U.S. policymaking to the creation of dozens of Behavioral Insights Units across the globe, understanding human behavior appears to have become a must for those seeking to devise policies that can effectively address important societal problems ranging from eliminating cheating on taxes to increasing school attendance, reducing smoking, fighting obesity, to increasing civic engagement. Our research contributes to a growing literature suggesting how behavioral science can inform policy. Specifically, our investigation implies that standard randomized controlled trials conducted to evaluate the efficacy of incentives with the goal of proving their policy value may overestimate their impact in natural environments because such trials invariably entail making the incentives on offer conspicuous, which does not always mirror natural contexts. We hope our work will inspire further research on inattention and how to best reduce it, ensuring that valuable policies achieve their well-intended goals of changing behavior for the better.

References

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science, 61*, 146-160. doi:10.1287/mnsc.2014.2091

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics, 95*(9), 1082-1095. doi:https://doi.org/10.1016/j.jpubeco.2011.03.003

Allcott, H., & Taubinsky, D. (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review, 105*(8), 2501-2538. doi: 10.1257/aer.20131564

Armstrong, M., & Vickers, J. (2012). Consumer protection and contingent charges. *Journal of Economic Literature, 50*, 477-493. doi:doi: 10.1257/jel.50.2.477

Asensio, O. I., & Delmas, M. A. (2015). Nonprice incentives and energy conservation. *Proceedings of the National Academy of Sciences, 112*(6), E510-E515. doi:10.1073/pnas.1401880112

Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica, 74*, 431-497. doi:doi:10.1111/j.1468-0262.2006.00668.x

Baca-Motes, K., Brown, A., Gneezy, A., Keenan, E. A., & Nelson, L. D. (2012). Commitment and behavior change: Evidence from the field. *Journal of Consumer Research*, 39(5), 1070-1084.

Bandura, A., Adams, N. E., & Beyer, J. (1977). Cognitive processes mediating behavioral change. *Journal of Personality and Social Psychology*, 35(3), 125.Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*,9(3), 226–232.

Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R block Fafsa experiment*. *The Quarterly Journal of Economics, 127*(3), 1205-1242. doi:10.1093/qje/qjs017

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5), 803-843.

Briss, P. A., Rodewald, L. E., Hinman, A. R., Shefer, A. M., Strikas, R. A., Bernier, R. R., . . . Williams, S. M. (2000). Reviews of evidence regarding interventions to improve vaccination coverage in children, adolescents, and adults. *American Journal of Preventive Medicine, 18*, 97-140. doi:https://doi.org/10.1016/S0749-3797(99)00118-X

Bronchetti, E. T., Huffman, D. B., & Magenheim, E. (2015). Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination. *Journal of Economic Behavior & Organization, 116*, 270-291. doi:https://doi.org/10.1016/j.jebo.2015.04.003

Bulman, G. B., & Hoxby, C. M. (2015). The returns to the federal tax credits for higher education. *Tax Policy and the Economy, 29*, 13-88. doi:10.1086/683364

Busse, M. R., Lacetera, N., Pope, D. G., Silva-Risso, J., & Sydnor, J. R. (2013). Estimating the effect of salience in wholesale and retail car markets. *American Economic Review, 103*, 575-579. doi:doi: 10.1257/aer.103.3.575

Chapman, G. B., Li, M., Colby, H., & Yoon, H. (2010). Opting in vs opting out of influenza vaccination. *JAMA, 304*, 43-44. doi:10.1001/jama.2010.892

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica, 77*(3), 909-931. doi:10.3982/ECTA7416

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review, 99*, 1145-1177. doi: 10.1257/aer.99.4.1145

Clippel, G. D., Eliaz, K., & Rozen, K. (2014). Competing for consumer inattention. *Journal of Political Economy, 122*(6), 1203-1234. doi:10.1086/677253

Dai, H., Mao, D., Volpp, K. G., Pearce, H. E., Relish, M. J., Lawnicki, V. F., & Milkman, K. L. (2017). The effect of interactive reminders on medication adherence: A randomized trial. *Preventive Medicine, 103*, 98-102. doi:https://doi.org/10.1016/j.ypmed.2017.07.019

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature, 47*(2), 315-372. doi: 10.1257/jel.47.2.315

Dohmen, T. J., Falk, A., Huffman, D., & Sunde, U. (2006). Seemingly irrelevant events affect economic perceptions and expectations: the FIFA World Cup 2006 as a natural experiment. doi:https://ssrn.com/abstract=928830

Ericson, K. M. (2017). On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation. *Journal of the European Economic Association, 15*, 692-719. doi:10.1093/jeea/jvw015

Federal Communications Commission (2015) Bill shock. Accessed September 29, 2016, https://www.fcc.gov/general/bill-shock

Finkelstein, A. (2009). E-ztax: Tax salience and tax rates*. *The Quarterly Journal of Economics, 124*(3), 969-1010. doi:10.1162/qjec.2009.124.3.969

Finkelstein, A., Linnan, L. A., Tate, D. F., & Birken, B. E. (2007). A pilot study testing the effect of different levels of financial incentives on weight loss among overweight employees. *Journal of Occupational and Environmental Medicine, 49*(9), 981-989. doi:10.1097/jom.0b013e31813c6dcb

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2011). Time discounting and time

    preference:. *Advances in Behavioral Economics,* 162-222. doi:10.2307/j.ctvcm4j8j.11

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify

    behavior. *Journal of Economic Perspectives, 25*(4), 191-210. doi: 10.1257/jep.25.4.191

Gneezy U., & Rustichini, A. (2000a). A fine is a price. *The Journal of Legal Studies, 29*(1), 1-17.

    doi:10.1086/46806

Gneezy, U., & Rustichini, A. (2000b). Pay enough or don't pay at all. *The Quarterly Journal of*

    Economics, 115(3), 791-810.Goldstein S. P., Goldstein, C. M., Bond, D. S., Raynor, H.

    A., Wing, R. R., & Thomas, J. G. (2019). Associations between self-monitoring and

    weight change in behavioral weight loss interventions. *Health Psychology, 38*(12), 1128-

    1136. doi:10.1037/hea0000800

Grubb, M. D. (2015). Consumer inattention and bill-shock regulation. *The Review of Economic*

    Studies, 82, 219-257. doi:10.1093/restud/rdu024

Hanna, R., Mullainathan, S., & Schwarzstein (2014). Learning through Noticing: Theory and

    Experimental Evidence in Farming, *Quarterly Journal of Economics,* 129(3), 1311-1353.

Heath, C. (1999). On the social psychology of Agency Relationships: Lay theories of Motivation

    Overemphasize extrinsic incentives. *Organizational Behavior and Human Decision*

    Processes, 78(1), 25-62. doi:10.1006/obhd.1999.2826

Hirvensalo, M., Telama, R., Schmidt, M. D., Tammelin, T. H., Yang, X., Magnussen, C. G., . . .

    Raitakari, O. T. (2011). Daily steps among Finnish adults: Variation by age, sex, and

    socioeconomic position. *Scandinavian Journal of Public Health, 39*(7), 669-677.

    doi:10.1177/1403494811420324

Imas, A. (2014). Working for the "warm glow": On the benefits and limits of prosocial

 incentives. *Journal of Public Economics, 114*, 14-18.

 doi:https://doi.org/10.1016/j.jpubeco.2013.11.006

Ito, K. (2015). Asymmetric incentives in subsidies: Evidence from a large-scale electricity rebate

 program. *American Economic Journal: Economic Policy, 7*, 209-237. doi:

 10.1257/pol.20130397

Jeffery, R. W. (2012). Financial incentives and weight control. *Preventive Medicine, 55*.

 doi:10.1016/j.ypmed.2011.12.024

John, L. K., Loewenstein, G., Troxel, A. B., Norton, L., Fassbender, J. E., & Volpp, K. G.

 (2011). Financial incentives for extended weight loss: A randomized, controlled trial.

 Journal of General Internal Medicine, 26(6), 621-626. doi:10.1007/s11606-010-1628-y

Kahneman D. (2011). *Thinking, Fast and Slow*. Farrar, Straus, and Giroux, New York, NY.

Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annual Review of*

 Economics, 4(1), 427-452.

Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2016). Getting to the top of mind:

 How reminders increase saving. *Management Science*, 62(12), 3393-3411.

Karlan, D., Ratan, A. L., & Zinman, J. (2014). Savings by and for the poor: A research review

 and agenda. *Review of Income and Wealth, 60*, 36-78. doi:10.1111/roiw.12101

Laibson, D. (2011). Golden eggs and hyperbolic discounting. *Advances in Behavioral*

 Economics, 429-457. doi:10.2307/j.ctvcm4j8j.20

Loewenstein, G., & Thaler, R. H. (1989). Anomalies: Intertemporal choice. *Journal of Economic*

 Perspectives, 3(4), 181-193. doi:10.1257/jep.3.4.181

Loewenstein, G., Asch, D. A., & Volpp, K. G. (2013). Behavioral economics holds potential to deliver better results for patients, insurers, and employers. *Health Affairs, 32*(7), 1244-1250. doi:10.1377/hlthaff.2012.1163

Long, M. (2004). The impact of asset-tested college financial aid on household savings. *Journal of Public Economics, 88*, 63-88. doi:https://doi.org/10.1016/S0047-2727(02)00136-6

Luca, M., & Smith, J. (2013). Salience in quality disclosure: Evidence from the U.S. news college rankings. *Journal of Economics & Management Strategy, 22*, 58-77. doi:10.1111/jems.12003

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic Theory* (Vol. 1): Oxford university press New York.

McClure, S. M., Daw, N. D., & Montague, P. R. (2003). A computational substrate for incentive salience. *Trends in neurosciences*, *26*(8), 423-428.

Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the united states, 2000. *JAMA*, 291(10), 1238-1245. doi:10.1001/jama.291.10.1238

O'Donoghue, T., & Rabin, M. (2011). Doing it now or later. *Advances in Behavioral Economics,* 223-251. doi:10.2307/j.ctvcm4j8j.12

Pope, D. G., & Pope, J. C. (2015). When Walmart comes to town: Always low housing prices? Always? *Journal of Urban Economics, 87*, 1-13. doi:https://doi.org/10.1016/j.jue.2014.10.004

Reimann, M., MacInnis, D., & Bechara, A. (2013). Reward Substitution: Incentivizing Consumers to Choose Smaller Portion Sizes, *NA - Advances in Consumer Research,* 41(1), eds. Simona Botti and Aparna Labroo, Duluth, MN : Association for Consumer Research.

Rogers, T., & Milkman, K. L. (2016). Reminders through association. *Psychological Science, 27*, 973-986. doi:10.1177/0956797616643071

Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics, 7*, 51-84. doi: 10.1257/app.20130327

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1*(1), 7-59. doi:10.1007/bf00055564

Schwartz, J., Mochon, D., Wyper, L., Maroba, J., Patel, D., & Ariely, D. (2014). Healthier by precommitment. *Psychological Science*, 25(2), 538-546.

Shafir, E. (2012). *The behavioral foundations of public policy*. Princeton, NJ: Princeton University Press.

Shea, S., DuMouchel, W., & Bahamonde, L. (1996). A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting. *Journal of the American Medical Informatics Association, 3*(6), 399-409. doi:10.1136/jamia.1996.97084513

Simon, H. A. (1947). *Administrative behavior: A study of the decision-making process in administrative organization*. New York: Macmillan.

Sindelar, J. L. (2008). Paying for performance: the power of incentives over habits. *Health Economics, 17*(4), 449-451. doi:10.1002/hec.1350

Stango, V., & Zinman, J. (2009). Exponential growth bias and household finance. *The Journal of Finance, 64*(6), 2807-2849. doi:10.1111/j.1540-6261.2009.01518.x

Szilagyi, P. G., Bordley, C., Vann, J. C., Chelminski, A., Kraus, R. M., Margolis, P. A., & Rodewald, L. E. (2000). Effect of patient reminder/recall interventions on immunization rates: A review. *JAMA*, 284(14), 1820-1827. doi:10.1001/jama.284.14.1820

Taubinsky, D., & Rees-Jones, A. (2017). Attention variation and welfare: Theory and evidence from a tax salience experiment. *The Review of Economic Studies, 85*(4), 2462-2496. doi:10.1093/restud/rdx069

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., & Staake, T. (2018). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management Science, 64*, 1458-1476. doi:10.1287/mnsc.2016.2646

Van den Steen, E. (2004). Rational overoptimism (and other biases). *American Economic Review, 94*(4), 1141-1151. doi:10.1257/0002828042002697

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: A randomized trial. *JAMA, 300*(22), 2631-2637. doi:10.1001/jama.2008.804

Volpp, K. G., Troxel, A. B., Pauly, M. V., Glick, H. A., Puig, A., Asch, D. A., . . . Audrain-McGovern, J. (2009). A randomized, controlled trial of financial incentives for smoking cessation. *New England Journal of Medicine, 360*(7), 699-709. doi:10.1056/NEJMsa0806819

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*: MIT press.

Zlatev J. J., Daniels, D. P., Kim, H., & Neale, M. A. (2017). Default neglect in attempts at social influence. *Proceedings of the National Academy of Sciences, 114*(52), 13643-13648. doi:10.1073/pnas.1712757114

Tables and Figures

Table 1: This table shows statistics from 3 weeks prior to the intervention up until the start of the intervention by experimental condition. Standard errors are in parentheses.

	(1)	(2)	(3)
	Conspicuous	Inconspicuous	(1) vs. (2), *p*-value
Mean Daily Steps Pre-Intervention	5,400.06	5,577.47	.15
	(89.40)	(85.21)	[*t*-test]
Fraction of People Who Recorded <2,000 Steps on Any Day Pre-Intervention	0.73 (0.01)	0.69 (0.01)	.07 [proportions test]
Monthly Email Open Rate Pre-Intervention	3.82	3.93	.36
	(2.51)	(2.73)	[*t*-test]
Kickoff Email Opened	0.64	N/A	N/A
	(0.02)		
Opened At Least One Notification Email	0.83	N/A	N/A
	(0.01)		
Observations	1,027	1,028	

Table 2: This table reports coefficient estimates from a series of ordinary least squares regression models (see Equation 1) predicting daily steps taken by a given user. Robust standard errors are clustered by participant and reported in parentheses. The analyzed data include observations of participants' daily steps from the 21 days before the intervention, 14 days during the intervention, and 21 days after the intervention. Model 1 is estimated using observations of all study participants (including participant-days with zero steps observed). Model 2 is estimated using observations of all participant-days with daily steps greater than 2,000. Model 3 replaces all daily step data below 2,000 steps, including missing observations, with an average of the individual's pre-intervention daily step counts stratified by day of week and excluding days with fewer than 2,000 steps.

	Model 1	**Model 2**	**Model 3**
(Conspicuous Incentive Condition) x (Treatment Period)	367.06***	229.04**	167.48**
	(105.71)	(83.63)	(64.76)
(Conspicuous Incentive Condition) x (0-2 Week Post-Treatment)	331.92**	210.93*	177.05*
	(119.61)	(92.60)	(71.03)
(Conspicuous Incentive Condition) x (3 Week Post-Treatment)	145.98	96.29	65.29
	(141.36)	(117.78)	(86.67)
Fixed effects for day of the year	Yes	Yes	Yes
Fixed effects for (user) x (day of the week)	Yes	Yes	Yes
Observations	115,080	85,789	115,080
R-squared	0.51	0.48	0.57
Clusters	2,055	2,029	2,055
Subset of Observations:	All Observations	Observations with > 2000 Daily Steps	All Observations, Imputing Observations With ≤ 2000 Steps

Standard errors in parentheses

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3: This table reports coefficient estimates from a series of ordinary least squares regression models (see Equation 1) predicting daily steps taken by a given user, estimated at each quartile of email open rate (number of emails opened per month) prior to the intervention. Robust standard errors are clustered by participant and reported in parentheses. The analyzed data include observations of participants' daily steps from the 21 days before the intervention, 14 days during the intervention, and 21 days after the intervention. These models are estimated using observations of all study participants (including participant-days with zero steps observed).

Number of emails opened per month	Daily Steps			
	≥ 1	≥ 1.61	≥ 2.86	≥ 4.29
	(Top 100%)	(Top 75%)	(Top 50%)	(Top 25%)
(Conspicuous Incentive Condition) x (Treatment Period)	367.06***	437.64***	511.50**	420.92**
	(105.71)	(114.29)	(64.76)	(191.96)
(Conspicuous Incentive Condition) x (0-2 Week Post-Treatment)	331.92**	352.28**	381.37*	122.30
	(119.61)	(127.54)	(153.96)	(209.62)
(Conspicuous Incentive Condition) x (3 Week Post-Treatment)	145.98	49.96	227.36	-30.36
	(141.36)	(150.03)	(175.72)	(233.29)
Observations	115,080	94,370	67,260	47,798
R-squared	0.51	0.52	0.53	0.53
Clusters	2,055	1,783	1,196	657

Standard errors in parentheses
$+ p < .10$, $* p < .05$, $** p < .01$, $*** p < .001$

Table 4. This table reports coefficient estimates from a series of ordinary least squares regression models (see Equation 2) predicting daily steps taken by a given user. Robust standard errors are clustered by participant and reported in parentheses. The analyzed data include observations of participants' daily steps from the 21 days before the intervention, 14 days during the intervention, and 21 days after the intervention, as well as the daily steps of a matched holdout group. Model 1 is estimated using observations of all study participants (including participant-days with zero steps observed). Model 2 is estimated using observations of all participant-days with daily steps greater than 2,000. Model 3 replaces all daily step data below 2,000 steps, including missing observations, with an average of the individual's pre-intervention daily step counts stratified by day of week and excluding days with fewer than 2,000 steps.

	Model 1	**Model 2**	**Model 3**
(In Experiment) x (Treatment Period)	-110.04	52.08	6.03
	(89.72)	(72.12)	(55.49)
(In Experiment) x (0-2 Weeks Post-Treatment Period)	-413.82**	-198.22	-260.71**
	(129.58)	(140.87)	(98.80)
(In Experiment) x (3 Weeks Post-Treatment Period)	-130.16	141.74	-26.54
	(123.58)	(101.51)	(76.15)
(Conspicuous Incentive Condition) x (Treatment Period)	406.42***	242.90**	188.87**
	(106.05)	(84.18)	(65.57)
(Conspicuous Incentive Condition) x (0-2 Weeks Post-Treatment Period)	337.04**	208.83*	176.54*
	(119.23)	(92.43)	(70.92)
(Conspicuous Incentive Condition) x (3 Weeks Post-Treatment Period)	145.98	95.00	65.29
	(141.04)	(117.67)	(86.47)
Wald Test ($\beta_4 - \beta_1$) *Difference in Coefficients*	516.46**	190.86	182.84+
	(174.17)	(138.55)	(107.10)
Wald Test ($\beta_5 - \beta_2$) *Difference in Coefficients*	750.85***	407.03*	437.26**
	(213.02)	(193.87)	(141.20)
Wald Test ($\beta_6 - \beta_3$) *Difference in Coefficients*	276.14	-46.76	91.83
	(234.35)	(194.61)	(144.17)
Observations	235,424	172,117	235,424

R-squared	0.33	0.26	0.32
Clusters	4,204	4,144	4,204
Subset of Observations:	All Observations	Observations with > 2000 Daily Steps	All Observations, Imputing Observations With ≤ 2000 Steps

Standard errors in parentheses

+ *p* < .10, * *p* <. 05, ** *p* < .01, *** *p* < .001

Figure 1: Plot of the difference in mean daily steps taken by participants in the conspicuous incentives condition and the inconspicuous incentives condition, using raw data from during and after the intervention.
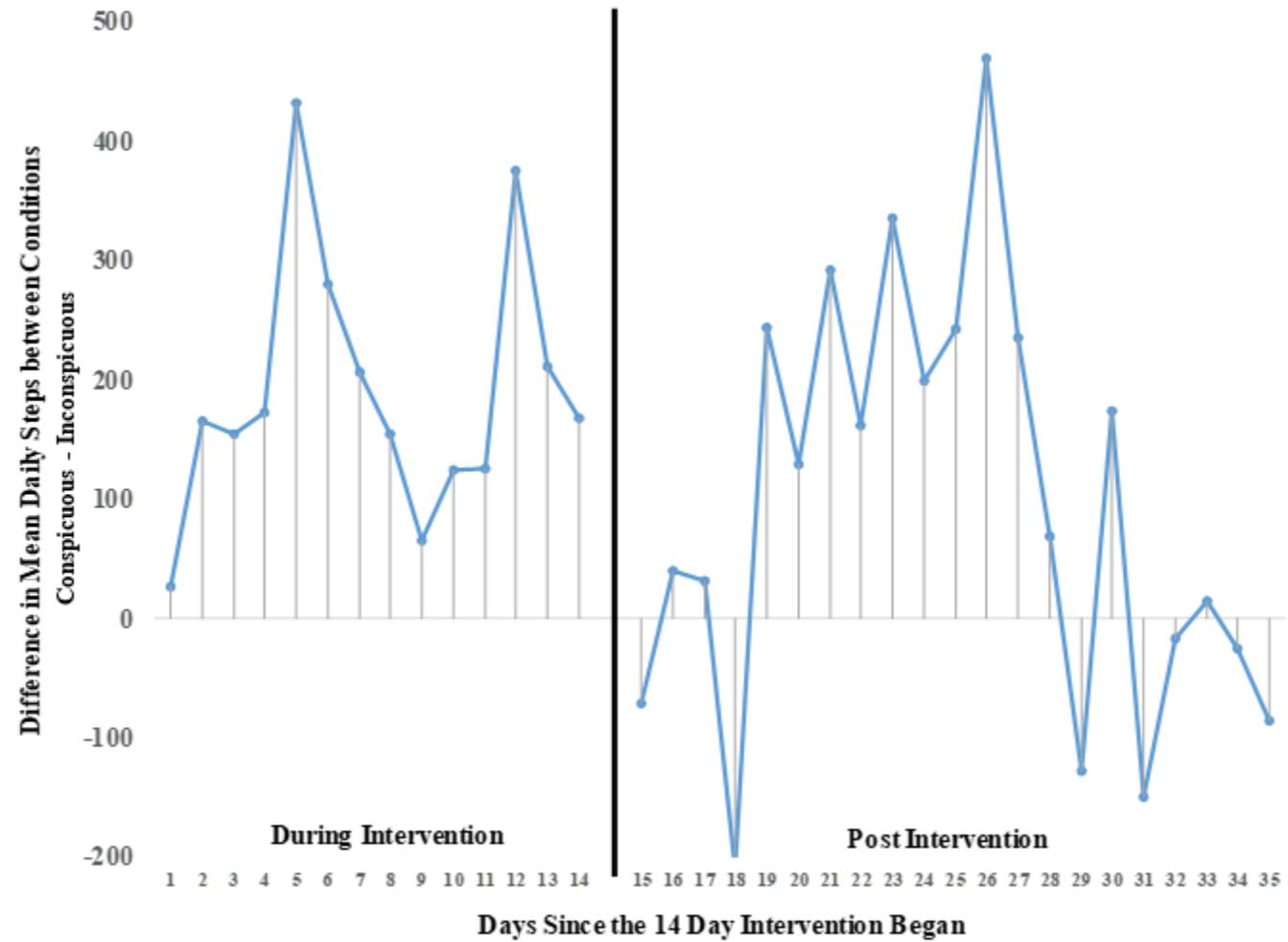
APPENDIX

Figure A1. Kickoff email.

Users in the conspicuous incentives condition received the following email on Jan. 26, 2015 (the day before incentives began):

Subject Line: New Program to Encourage You to Walk (earn Bonus Points)
Message:

AchieveMint

⬆Bonus Points Opportunity

New Program to Encourage You to Walk (earn Bonus Points)

Tomorrow is the first day of a two week walking program designed in partnership with experts at Harvard and the University of Pennsylvania to get you moving. Tomorrow and every day after that for the next two weeks, we'll encourage you to walk by multiplying the points you earn for walking by 40.

To push you to walk more, your bonuses from Achievemint over the next two weeks will follow this schedule:

JANUARY 2015

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27 **40x**	28 **40x**	29 **40x**	30 **40x**	31 **40x**

FEBRUARY 2015

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1 **40x**	2 **40x**	3 **40x**	4 **40x**	5 **40x**	6 **40x**	7 **40x**
8 **40x**	9 **40x**	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28

We will be sending you reminders every two days about upcoming bonuses.

We hope that this program will help you improve your walking habits!
For the next two weeks, we'll be emailing you every other day about these bonuses.
If you don't want to receive these emails, please click here.

Figure A2. Notification emails.

Users in the conspicuous incentives condition received the following notification email every other day (day 1, day 3,…, day 13) during the 14-day incentive program.

Subject Line: Program to Increase Walking (Earn Bonus Points Today and Tomorrow)
Message:

⬆Bonus Points Opportunity

Program to Increase Walking (Earn Bonus Points Today and Tomorrow)

You are in the middle of a two week walking program designed in partnership with experts at Harvard and the University of Pennsylvania to get you moving. Today and tomorrow, we'll encourage you to walk: all points you earn for walking will be multiplied by 40.

To push you to walk more, your bonuses from Achievemint over the next two weeks will follow this schedule:

For the duration of this two week period, we'll be emailing you every other day about these bonuses. If you don't want to receive these emails, please click[here](#).

Figure A3. Experiment 2 Conspicuousness Manipulation.

Inconspicuous Condition

During this study, you will answer several questions about yourself and your opinions. It is important that you share your true opinions as there are no right or wrong answers.
In Part I of this study, you will share your opinion about a picture.
In Part II of this study, you will answer some survey questions. During Part II of this study, you will have the opportunity to earn an additional bonus. Four of the multiple choice questions will include an option to select 'other' and write in an open-ended response. You can earn a four cent bonus for each of these questions by selecting the 'other' multiple choice answer and writing in the phrase "seagulls fly over the sea." If you wish to write in your own response to this question, you may write it after this bonus phrase.

Please click >> to proceed to Part I.

Conspicuous Condition

During this study, you will answer several questions about yourself and your opinions. It is important that you share your true opinions as there are no right or wrong answers.

In Part I of this study, you will share your opinion about a picture.

In Part II of this study, you will answer some survey questions. During Part II of this study, you will have the opportunity to earn an additional bonus. Four of the multiple choice questions will include an option to select 'other' and write in an open-ended response. **You can earn a four cent bonus for each of these questions by selecting the 'other' multiple choice answer and writing in the phrase "seagulls fly over the sea."** If you wish to write in your own response to this question, you may write it after this bonus phrase.

Please click >> to proceed to Part I.