Is Attention Produced Optimally?

Erin T. Bronchetti, Judd B. Kessler, Ellen B. Magenheim, Dmitry Taubinsky, and Eric Zwick

December 2021

PRC WP2022-08
Pension Research Council Working Paper
Pension Research Council
The Wharton School, University of Pennsylvania
3620 Locust Walk, 3302 SH-DH
Philadelphia, PA 19104-6302
Tel.: 215.573.3414 Fax: 215.573.3418
Email: prc@wharton.upenn.edu
http://www.pensionresearchcouncil.org

Is Attention Produced Optimally?

Abstract

This paper investigates whether people know their attention cost functions and allocate attention optimally. We characterize how demand for instruments that reduce attention costs varies with the pecuniary incentives to be attentive, under the null hypothesis of correct perceptions and optimal choice. We examine if the optimality conditions are satisfied in three experiments. The first is a field experiment (n = 1373) with an online education platform, in which we randomize incentives to complete course modules and incentives to utilize a plan-making tool to complete the modules. In the second experiment (n = 2306), participants must complete a survey in the future. We randomize survey-completion incentives and how long participants must wait to complete the survey, and we elicit willingness to pay for reminders. The third experiment (n = 1465) involves a psychometric task in which participants must identify whether there are more correct or incorrect mathematical equations in an image. We vary incentives for accuracy, elicit willingness to pay to reduce task difficulty, and examine the impact of learning and feedback. In all experiments, demand for reducing attention costs increases as incentives for accurate task completion increase. However, in all experiments—and across all conditions—our tests imply that this increase in demand is too small relative to the null of correct perceptions. These results suggest that people may be uncertain or systematically biased about their attention cost functions, and that experience and feedback do not necessarily eliminate mistakes.

Keywords: attention, attention cost, optimality conditions

Erin T. Bronchetti
Swarthmore College
500 College Avenue
Swarthmore, PA 19081
ebronch1@swarthmore.edu

Judd B. Kessler
University of Pennsylvania and NBER
320 Vance Hall
3733 Spruce Street
Philadelphia, PA 19104
judd.kessler@wharton.upenn.edu

Ellen B. Magenheim
Swarthmore College
500 College Avenue
Swarthmore, PA 19081
emagenh1@swarthmore.edu

Dmitry Taubinsky
UC Berkley and NBER
530 Evans Hall #3880
Berkeley CA 94720
dmitry.taubinsky@berkeley.edu

Eric Zwick
University of Chicago Booth and NBER
5807 South Woodlawn Avenue
Harper Center HC431
Chicago, IL 60637
ezwick@chicagobooth.edu

Is Attention Produced Optimally?*

Erin T. Bronchetti, Judd B. Kessler, Ellen B. Magenheim
Dmitry Taubinsky, Eric Zwick

December, 2021

Abstract

This paper investigates whether people know their attention cost functions and allocate attention optimally. We characterize how demand for instruments that reduce attention costs varies with the pecuniary incentives to be attentive, under the null hypothesis of correct perceptions and optimal choice. We examine if the optimality conditions are satisfied in three experiments. The first is a field experiment ($n = 1373$) with an online education platform, in which we randomize incentives to complete course modules and incentives to utilize a plan-making tool to complete the modules. In the second experiment ($n = 2306$), participants must complete a survey in the future. We randomize survey-completion incentives and how long participants must wait to complete the survey, and we elicit willingness to pay for reminders. The third experiment ($n = 1465$) involves a psychometric task in which participants must identify whether there are more correct or incorrect mathematical equations in an image. We vary incentives for accuracy, elicit willingness to pay to reduce task difficulty, and examine the impact of learning and feedback. In all experiments, demand for reducing attention costs increases as incentives for accurate task completion increase. However, in all experiments—and across all conditions—our tests imply that this increase in demand is too small relative to the null of correct perceptions. These results suggest that people may be uncertain or systematically biased about their attention cost functions, and that experience and feedback do not necessarily eliminate mistakes.

*Bronchetti and Magenheim: Swarthmore College. Kessler: University of Pennsylvania and NBER. Taubinsky: UC Berkeley and NBER. Zwick: University of Chicago Booth and NBER. We thank Andrew Caplin, Mark Dean, Xavier Gabaix, Stephen O'Connell, Devin Pope, three grant reviewers at the Russell Sage Foundation, and seminar and conference participants for helpful comments and advice. We thank Alexander Hirsch, Stephanie Nam, Laila Voss, and Caleb Wroblewski for excellent research assistance. We gratefully acknowledge Mike Walmsley and CodeAvengers.com for their support with the education experiment. We gratefully acknowledge research funding from the Russell Sage Foundation, Swarthmore College, the Boettner Center, the Wharton School, the Wharton Behavioral Lab, and the Alfred P. Sloan Foundation. The first experiment was approved by the Swarthmore (covering Haverford and Muhlenberg), Bryn Mawr, Lafayette, and Ursinus IRBs, numbers: 14-15-065, R17-042, AY1617-12, 01-18-17. The second and third experiments were approved by the University of Pennsylvania IRB, number 832335. The opinions expressed in this paper are solely the authors', and do not necessarily reflect the views of any individual or institution listed above.

A large and rapidly growing body of work in economics and cognitive science shows that attention is costly and at least partly controlled (for recent reviews in economics, see Caplin, 2016; Maćkowiak et al., forthcoming; Gabaix, 2019). There is a growing recognition that, like many other types of costly effort decisions that economists have analyzed for decades, mental effort is also costly and deliberately deployed (Shenhav et al., 2017). A powerful modeling approach taken by the *rational inattention* literature is to assume that individuals have complete information about their attention cost functions and that they choose their attention strategies optimally. These assumptions are worth testing, as people appear to have incomplete knowledge of their production functions in many settings. For example, individuals may be uncertain about the returns to human capital investments (Wiswall and Zafar, 2018, 2021), may systematically overestimate their productivity (Hoffman and Burks, 2020), or may be influenced by behavioral biases in real-effort tasks (DellaVigna and Pope, 2017).

In this paper, we investigate how people perceive their attention cost functions and whether they use this knowledge to allocate attention optimally. We first develop a methodology for measurement. We then deploy this methodology in three complementary experiments that study whether people properly value opportunities that reduce demands on their *mental bandwidth* (i.e., reduce the cost of being attentive). The first two experiments focus on people's ability to remain attentive to a future task. The third experiment explores an attention-demanding psychometric task, as in recent experiments studying rational inattention models (Dean and Neligh, 2018; Caplin et al., 2020; Ambuehl et al., 2020). To fix terminology, we use the term bandwidth enhancements (BEs) to refer to the plan-making and reminder tools in the first two experiments and to the reductions in required mental effort in the third experiment.

Our theoretical approach builds on the Caplin et al. (2020) characterization of costly attention models with a competitive supply framework, and the insights on bounded rationality and Slutsky symmetry introduced by Gabaix (2014). Our approach clarifies the difficulty with assessing whether people optimally value attention-cost reductions just by examining willingness to pay for BEs and how BEs affect behavior. The intuition is that individuals may particularly like or dislike a given BE for reasons unrelated to its impact on behavior. The main idea of our approach is to examine how pecuniary incentives to complete a task affect demand for BEs that aid task completion. More specifically, the first prediction from our model is a precise condition on how willingness to pay for BEs changes with the pecuniary rewards for the task, under the null of full optimality and correct perceptions. The second prediction is a form of a Slutsky symmetry condition, which

states that the impact of task-completion incentives on take-up of the BE is equal to the impact of the price of the BE on the propensity to complete the task.[1]

Guided by this framework, we carry out three experiments. The first experiment was run in the field with 1373 students and alumni from six Philadelphia-area colleges who enrolled in an 8-week online coding course. The experiment randomized incentives to complete three 15-minute coding lessons each week and randomized incentives to make a plan to complete three 15-minute course modules each week. Making a plan involved clicking a link that automatically created three 15-minute events in the participant's online calendar of choice for the following week and allowed the participant to rearrange the planned events to suit their schedule.

We document three key facts in our first experiment. First, use of our plan-making tool increased the likelihood of completing coding lessons, especially in the initial weeks. Second, take-up of our plan-making tool was elastic to the direct incentives for plan-making, but remained below 100 percent, even with the incentives. The combination of incomplete take-up and the positive elasticity suggests that the use of our tool imposes internal or "nuisance" costs on at least some individuals. Third, we find that take-up of our plan-making tool increased with incentives for *completing coding lessons*.

This third result is consistent with the *qualitative* rational inattention prediction that individuals should value BEs more as the rewards for completing a task increase. At the same time, we estimate that the impact of completion incentives on plan-making is quantitatively too small relative to the Slutsky symmetry condition, suggesting that participants undervalue the plan-making tool. However, our confidence intervals are wide and do not permit us to reject the null of full optimality under correct perceptions.

Our second experiment is an online survey-completion experiment that elicits richer data that allows us to fully quantify the demand for a reminder technology and to test the first prediction from our model. The study was conducted on Amazon Mechanical Turk (MTurk) with 2306 participants. Study participants were offered a bonus (either \$3, \$4, \$11, or \$12) for completing a survey that would only be accessible for a week-long period after a delay (either 2 days, 1 week, 3 weeks, or 6 weeks). Prior to randomizing participants into one of the four possible delays and one of the four possible survey-completion rewards, we elicited participants' willingness to pay (WTP) for a set of three reminder emails for each possible delay-reward pair. Our procedure also generated exogenous varia-

[1] See Gabaix (2019) for a discussion about exploiting Slutsky symmetry as an empirical strategy for testing limited attention. See Gabaix (2014) and Abaluck and Adams-Prassl (2021) for an implementation of such a test in the context of misperceived product attributes.

tion in whether participants actually received the reminder emails, allowing us to estimate the effect of reminder emails on survey completion.

We find that survey completion increased with incentives and decreased with delay, while the impact of reminders decreased with incentives and increased with delay. The average impact of reminders on completing the survey was 29 and 16 percentage points for *low* (\$3 or \$4) and *high* (\$11 or \$12) task-completion incentives, respectively. Across the eight different possible delay-reward pairs, the impact of reminders ranged from -7 percentage points ($se = 6.00$) to 40 percentage points ($se = 5.88$).

This set of findings informs several hypotheses about attention in this type of setting. First, the negative effect of delay on task completion and the positive effect of delay on the impact of reminders is consistent with the attention/memory decay curves proposed by cognitive psychologists (see, e.g., Mullainathan, 2002; Ericson, 2017, for reviews). Second, the fact that, at high incentives, task completion was higher but the impact of reminders was lower is consistent with individuals exerting more costly attention to keep the task top of mind when stakes are higher.

Our second set of findings from our second experiment is that, while WTP for reminders increased with the size of the bonus for survey completion, the increase was too small relative to the null of correct perceptions of attention costs. The theory implies that under this null, an extra \$1 of task-completion incentives should increase WTP by \$0.29 and \$0.16 in the *low* and *high* incentive groups, respectively. Instead, WTP increased by \$0.07 ($se = 0.017$) and \$0.02 ($se = 0.047$) in those two groups. Using the variation in WTP and the effects of reminders across all eight conditions, we estimate a model of how people's perceived effects of reminders vary with the actual effects. We find that people uniformly underestimate the effects of reminders by 84 percent, rather than underestimating the effects when they are small and overestimating them when they are large.

Our third experiment expands the scope of our analysis in two ways. First, it involves a different domain of decision-making, illustrating the breadth and portability of our methods. Second, it allows us to study how learning and feedback affect people's perceptions of the value of BEs. Specifically, this experiment involved a series of tasks, first utilized by Ambuehl et al. (2020), in which participants were shown an image with a set of arithmetic equations that were either correct (e.g., $10+12=22$) or incorrect (e.g., $10+12=23$). Participants were asked to indicate whether an image contained more correct or more incorrect equations, and were rewarded for accuracy. All participants completed baseline tasks, in which the image had 100 equations and either 60 or 40 percent of equations were correct.

We randomly assigned participants to also do one of two less cognitively demanding

versions of the task. In the *length* condition, the easier task (i.e., the BE in this setting) involved only 10 equations. In the *discernibility* condition, the easier task had either 95 or 5 correct equations. Participants completed two blocks of seven tasks, with each block containing three baseline tasks, three easier tasks, and one task that might be affected by the participants' preferences. Analogous to the second experiment, we varied incentives for task completion (i.e., accuracy in this case) and, prior to each block, we elicited individuals' WTP to make the remaining task easier for the different incentive levels. Additionally, prior to the WTP elicitation in the second block, we randomly gave some participants feedback about their performance in the first block on the hard and easy tasks.

We find that participants were 19 ($se = 0.647$) and 26 ($se = 0.607$) percentage points more likely to accurately answer the easier tasks in the length and discernibility conditions, respectively, and these differences were nearly identical across both blocks. Under the null of correct perceptions, participants' WTP to decrease difficulty in the length and discernibility conditions should thus increase by \$0.19 and \$0.26 with each extra dollar of accuracy incentives. However, we find that in the first block, the WTP increases are only \$0.10 ($se = 0.034$) and \$$-0.01$ ($se = 0.030$) in the length and discernibility conditions, respectively. Thus, as in the other experiments, participants undervalue the BEs.

Our data from this experiment provide two additional lessons that complement the insights from the first two experiments. First, in the first block, participants' valuations of the discernibility BE are significantly more biased than their valuations of the length BE. This fact illustrates the potential context-specificity of BE valuations, and the need for portable methods that can quantify them across contexts. Second, we study the effects of learning and feedback. We find that in the second block, participants' WTP to make the task easier increases by only \$0.03 ($se = 0.025$) and \$0.03 ($se = 0.024$) for each extra dollar of accuracy incentives in the length and discernibility conditions, respectively. In the length condition, this result reflects a significant *decrease* relative to the first block. We show that this decrease is concentrated among people who received feedback that they performed at least as well in the longer version of the task than the shorter version, which suggests that—on average—people overweighted experiences in which the length-decreasing BE did not improve performance. This pattern is consistent with recent work on mis-specified learning (e.g., Heidhues et al., 2018; Gagnon-Bartsch et al., 2021), which suggests that experience and feedback do not necessarily eliminate mistakes.

Our results contribute to the literature in several ways. First, we build on the supply theory framework developed by Caplin et al. (2020) to develop a method to assess whether individuals understand their attention production functions and to test models of rational

inattention. Despite the recent proliferation of work on rational inattention, surprisingly little work has been done on individuals' understanding of the limitations of their attention.[2] Our experiments illustrate how our method can be applied both to the kinds of psychometric settings where rational inattention models have traditionally been tested, as well as to settings concerning behaviors such as education and health investments where the study of attention has been more reduced-form.

Second, a large body of work looks at the impact of BEs, such as planning prompts and reminders on behaviors such as medical compliance, educational attainment, savings, loan repayment, wage reporting, voting, and charitable donation.[3] We advance this literature by studying individuals' demand for BEs. Our approach sheds light on whether provision of such BEs is efficient. If individuals valued these BEs optimally, then external provision of the BEs would be inefficient because the market already provides individuals with many opportunities to acquire reminder technologies and plan-making tools in the form of various smartphone and computer applications, online calendars, smart caps on pill bottles, and so on. As noted above, incomplete take-up of BEs does not by itself imply that people undervalue them, because in addition to any pecuniary costs, reminders and plan-making tools may carry private nuisance costs (see, e.g., Damgaard and Gravert, 2018), time costs, or detract scarce attention from other important tasks (Nafziger, 2020).

Closest to our second experiment, Ericson (2011) and Tasoff and Letzler (2014) conduct lab experiments that find that individuals' willingness to pay for a rebate exceeds the expected returns because individuals' use of the rebate is low. Their results suggest overestimation of future attention to the rebate, although other biases, such as Tasoff and Letzler's (2014) proposed *weak cost-salience*, plausibly also play a role. Our approach and results from the second experiment complement Ericson (2011) and Tasoff and Letzler (2014) in a few ways. First, overconfidence about one's baseline level of attention need not imply under-appreciation of the incremental impact of BEs, and vice versa. Second, by directly estimating individuals' (mis)valuations of BEs, our method allows us to directly speak to how much take-up of BEs should be encouraged through subsidies or other interventions. Third, the richness of our second experiment provides new insights about variation in at-

[2]There is more work on the optimality of individuals' information acquisition strategies. See, e.g., Gabaix et al. (2006); Hanna et al. (2014); Bartoš et al. (2016); Martin (2016); Dean and Neligh (2018); Ambuehl et al. (2020); Caplin et al. (2020); Carvalho and Silverman (2019). See also Altmann et al. (forthcoming) on costly attention and spillover effects.

[3]See, e.g., Nickerson and Rogers (2010); Milkman et al. (2011); Altmann and Traxler (2014); Castleman and Page (2016); Bronchetti et al. (2015); Karlan et al. (2016); Calzolari and Nardotto (2017); Damgaard and Gravert (2018); Marx and Turner (2019); Zhang et al. (2021). See also Carrera et al. (2018) and Oreopoulos et al. (forthcoming) for examples of null effects.

tention, such as our result that people are more attentive at higher stakes.

Finally, our work relates to the broader literature that studies whether individuals' beliefs are well-calibrated. A common approach is to directly elicit individuals' beliefs. However, the beliefs that individuals state in an abstract elicitation are not necessarily the decision weights that individuals apply in all real-stakes decisions because of salience and context effects (see, e.g., Bernheim and Taubinsky, 2018 for a recent discussion). For example, many individuals know how large sales taxes are and what products they apply to, but still neglect to incorporate them into their decisions (Chetty et al., 2009; Taubinsky and Rees-Jones, 2018). Similarly, individuals might have an abstract understanding of how BEs affect behavior yet still undervalue them in real-stakes decisions. Our approach is thus a useful complement to this other work.

The paper proceeds as follows. Section 1 presents our theoretical framework. Sections 2-4 present the designs and results from our three experiments. Section 5 concludes.

1 Theoretical Framework

1.1 Model

We consider individuals who choose a level of costly attention, which is needed to (correctly) complete a task. The level of attention can correspond to the likelihood of being attentive to the task in the future, as in our first two experiments, or to the likelihood of correctly solving a cognitively demanding task, as in our third experiment. Our modeling of attention as a production technology uses results in Caplin et al. (2020), which shows that standard rational inattention models (e.g., Sims, 2003; Matějka and McKay, 2015; Caplin and Dean, 2015; Caplin et al., 2019) can be represented by a production model in which individuals pay a cost to obtain a probability μ of taking the right action. The Caplin et al. (2020) results imply that our framework makes minimal assumptions about the structure and dimensionality of attention allocation.

We begin by considering a setting where individuals know their cost functions without uncertainty, as in the prior theoretical work. Formally, individual i first makes a choice $j \in \{0, 1\}$ between attention cost functions $K_i^0(\mu)$ and $K_i^1(\mu)$, where μ is the probability of being attentive to the task. We think of K_i^1 as constituting a bandwidth enhancement (BE) over K_i^0. In the context of our first two experiments, this is a choice of whether individuals take up our plan-making or reminder tools. In the context of our third experiment, this is a choice of whether to make the task less cognitively demanding. We let p denote the

incremental cost of choosing $j = 1$ over $j = 0$. In our first experiment, $-p$ corresponds to the incentives we create for choosing our plan-making tool, while in our second and third experiments, p is the price of reminders or making the task easier.[4]

After choosing $j \in \{0, 1\}$, individuals choose their level of attention $\mu \in [0, 1]$, at cost $K^j(\mu)$. In our first two experiments, this is the likelihood of having the task top of mind. This choice could involve other ways of increasing attention to the task, such as setting their own reminders, engaging in internal "rehearsal" (e.g., Mullainathan, 2002), or asking others to remind them. In our third experiment, this is the likelihood of correctly identifying the state, which is a function of the amount of mental effort (including time) that individuals devote to learning the true state of the world in the task.[5] Given an attention technology j, the net utility benefit of an attention level μ is

$$\mu B_i(r) - K_i^j(\mu),$$

where $B_i(r)$ scales the benefits of being attentive given a task-completion or accuracy reward r. The microfoundations of $B_i(r)$ differ across settings and are as follows in our three experiments.

In the third experiment, $B_i(r) = r$ is the financial payoff from correctly identifying whether there are more correct or incorrect equations in a task. In this setting, we let $a = 1$ denote the event in which the individual correctly solves the task, and we let $a = 0$ denote the event in which the individual solves it incorrectly. This formalism can also apply to settings where employees exert mental effort under a piece-rate incentive scheme (e.g., Dean, 2019; Kaur et al., 2021; Bessone et al., forthcoming), where the choice of K^j can capture the choice of task difficulty, decision aids, or the level of distraction in the environment.

In our first two experiments, we let $a = 1$ denote task completion and we let $a = 0$ denote not completing the task. We assume that if individuals are inattentive to the task, they default to $a = 0$. If individuals are attentive, they choose whether or not to complete the task. The net-of-cost benefits of completing the task are given by $b_i + r$, where r is the financial payoff of completing the task and b_i is distributed according to a differentiable

[4]Individual differences in K_i^0 and K_i^1 could result from individual differences in baseline attentiveness; differences in how well-suited the BE is to an individual's needs; differences in the nuisance costs of reminders, and the personal and social costs of failing to execute a plan that one creates; or (in reduced-form) differences in the indirect costs of having one's attention to other activities reduced.

[5]In principle, we could model the first- and second-step choices as occurring simultaneously. However, because the second-step choices are unobservable to the analyst, we formally distinguish them from the observable choices of attention technologies in our experiments.

density function g_i and is realized after individuals choose their attention strategies K^j and μ. Conditional on being attentive, individuals thus choose $a = 1$ if and only if $b_i + r > 0$, and thus,

$$B_i(r) := \int_{b+r>0} (x+r) g_i(b) db. \tag{1}$$

Under the assumption that utility is locally linear in the pecuniary incentives, the optimal choice of j and μ maximizes $(b_i + r)\mu - K_i^j(\mu) - pj$. Our main result characterizes testable restrictions of the optimality assumption on a set of statistics that we measure in our two experiments. The first statistic is the willingness to pay (WTP) for technology $j = 1$; that is, the highest p at which $j = 1$ is preferred to $j = 0$. Note that if the nuisance cost of $j = 1$ is sufficiently high, this statistic can be negative, even if $j = 1$ lowers the marginal cost of attention. Average WTP is given by

$$\bar{W}(r) := \mathbb{E}\left[\max_{\mu}\left((b_i + r)\mu - K_i^1(\mu)\right) - \max_{\mu}\left((b_i + r)\mu - K_i^0(\mu)\right)\right].$$

We measure WTP directly in our second and third experiments. Our results below also involve the following statistics, which we measure in our experiments: $Pr(j = 1|p, r)$, the probability of individuals choosing technology $j = 1$ given financial incentives p and r; $Pr(a = 1|p, r)$, the probability of individuals completing the task (i.e., choosing $a = 1$) given incentives p and r; $Pr(a = 1|j, r)$, the probability of individuals choosing $a = 1$ if exogenously assigned attention technology j.

1.2 Main Results

Our main assumption—which we state formally in Appendix A.1—is that individual differences are sufficiently "smoothly distributed" such that $\bar{W}(r)$, $Pr(j = 1|p, r)$, and $Pr(a = 1|p, r)$ are differentiable functions of p and r. Under this assumption, optimal attention choice implies the testable restrictions below. Appendix A.4 provides an instructive graphical argument that sketches ideas behind our main results.

Proposition 1. *Define*

$$D(a = 1|r) := Pr(a = 1|j = 1, r) - Pr(a = 1|j = 0, r).$$

Average willingness to pay for the attention-increasing technology, as a function of the task-completion incentive r, satisfies

$$\frac{d}{dr}\bar{W}(r) = D(a=1|r). \tag{2}$$

The likelihood of choosing technology $j=1$ and the likelihood of completing the task, as functions of the task-completion incentive r and the technology price p, satisfy the equality

$$\frac{d}{dr}Pr(j=1|p,r) = -\frac{d}{dp}Pr(a=1|p,r) \tag{3}$$

$$= -\frac{d}{dp}Pr(j=1|p,r)D(a=1|r). \tag{4}$$

Equation (2) of Proposition 1 states that, if technology $j=1$ increases individuals' likelihood of choosing $a=1$ by, e.g., 10 percentage points under incentive r, then a small increase dr in r should increase individuals' average willingness to pay for $j=1$ by approximately $dr \times 0.1$. The result and intuition follow from the Envelope Theorem. Optimality implies that a small increase dr in the task incentive should be worth $Pr(a=1|j=1,r)dr$ to individuals exogenously assigned technology $j=1$, and should be worth $Pr(a=1|j=0,r)dr$ to individuals exogenously assigned technology $j=0$. Consequently, the average impact on WTP for $j=1$ is $D(a=1|r)dr$.

The first condition in Proposition 1 is a limit result in the sense that it applies to marginal changes in the task-completion incentive. The corollary below clarifies how this condition can be used to evaluate "small" but not "vanishing" changes.

Corollary 1. *Suppose that terms of order $\frac{d}{dr}D(a=1|r)(\Delta^2)$ are negligible. Then*

$$\bar{W}(r+\Delta) - \bar{W}(r) \approx \frac{\Delta}{2} \cdot \left(D(a=1|r) + D(a=1|r+\Delta)\right). \tag{5}$$

In words, the approximation in equation (2) is valid as long as the impact of the BE on behavior changes negligibly with respect to a small change in the incentive, which we show is the case in our second and third experiments. Note that dividing both sides by Δ and taking the limit as $\Delta \to 0$ in equation (5) produces the exact equality in Proposition 1.[6]

The condition in equation (2) requires rich data that is difficult to collect in some field settings and that we do not have in our first experiment. Equation (3) builds on equation (2) by characterizing how the probability of choosing $j=1$ and the probability of choosing $a=1$ are related to each other. The condition in equation (3) formalizes the basic intuition that

[6]Note that Δ need not necessarily be small. What is important is that $D(a=1|r+\Delta) - D(a=1|r)$ is small. For example, if this statistic is zero, then the condition in Corollary 1 is exact even for large Δ.

if attention is allocated optimally, then increasing the incentives for choosing $a = 1$ should increase the desire to adopt a technology that increases the likelihood of choosing $a = 1$. But while the qualitative comparative static could still be consistent with individuals under- or overvaluing the benefits of BEs, the quantitative condition clarifies exactly how much individuals should seek attention technology improvements. The condition in equation (4) is a restatement of the condition in (3) that follows from algebraic manipulations of the expression for $Pr(a = 1|p, r)$.

The condition is a variation on the Slutsky symmetry condition that cross-price elasticities of compensated demand functions must be equal to each other, and is analogous to the tests of sparse-max decision making derived in Gabaix (2014). Intuitively, $-\frac{d}{dp}Pr(a = 1|p, r)$ is an indication of how adoption of technology $j = 1$ affects the probability of choosing $a = 1$. In our online education experiment, this derivative is the average impact of our plan-making incentives on the likelihood of completing course modules. The higher this number is, the higher the impact of our plan-making tool on the likelihood of completing the course modules will be. And the higher the impact of the plan-making tool, the higher the impact of a small change in r on its value, as formalized in the first part of Proposition 1. This translates to a higher derivative $\frac{d}{dr}Pr(j = 1|p, r)$.[7]

More generally, the statistics in Proposition 1 reveal individuals' subjective perceptions of the value of the BE.

Proposition 2. *Let $\tilde{D}_i(a = 1|r)$ denote person i's perception of the effect of the BE on the likelihood of choosing $a = 1$, at incentive r. Let $\mathbb{E}_i[\tilde{D}_i(a = 1|r)]$ denote the average of these perceptions. Then*

$$\frac{d}{dr}\bar{W}(r) = \mathbb{E}_i[\tilde{D}_i(a = 1|r)] \tag{6}$$

$$\frac{d}{dr}Pr(j = 1|p, r) = -\frac{d}{dp}\mathbb{E}_i\left[Pr_i(j = 1|p, r)\tilde{D}_i(a = 1|r)\right]. \tag{7}$$

[7]Note that any data set that can be used to test condition (2) can be used to test the Slutsky symmetry condition as well. To see this, first note that eliciting individuals' WTP for $j = 1$ at incentive r is equivalent to eliciting the demand curve for $j = 1$ at incentive r, which means that this data set identifies $Pr(j = 1|p, r)$ for all p and for each task-completion incentive r utilized in the experiment. Thus, $\frac{d}{dr}Pr(j = 1|p, r)$ and $\frac{d}{dp}Pr(j = 1|p, r)$ are identified in this data set. Second, the right-hand-side term of (2), $D(a = 1|r)$, is identified by assumption. Thus, all of the statistics necessary to test (4), and therefore also (3), are available. Note, however, that the Slutsky symmetry conditions use strictly less data than condition (2): these conditions consider the function $Pr(j = 1|p, r)$ only in a neighborhood around a single incentive level p, while condition (2) considers $Pr(j = 1|p, r)$ across all possible values p for which $Pr(j = 1|p, r) \in (0, 1)$.

Intuitively, the result in Proposition 1 characterizes the predictions $\frac{d}{dr}\bar{W}_i(r)$ and $\frac{d}{dr}Pr_i(j = 1|p,r)$ that an analyst with beliefs $\tilde{D}_i(a = 1|r)$ would make. Taking expectations over i gives the result.

1.2.1 Remarks and Qualifications

Differences in fixed costs, $K_i^1(0) - K_i^0(0)$, may result from the potential nuisance costs of attention-improving technologies, which is consistent with negative WTP for reminders by some individuals in our second experiment. Thus, the value of a BE cannot be equated with its impact on the change in expected earnings, $rPr(a = 1)$. Simply documenting that, for example, individuals' valuations for a reminder that increases their chance of earning \$10 by 10% is smaller than \$1 is not a rejection of correct valuation of the reminder, because nuisance costs could decrease the value of the reminder. Our more robust test focuses instead on how individuals' valuations of the BE change as the pecuniary incentives for being attentive change.

It is also helpful to note that condition (2) is a test of whether individuals correctly value the BE *on the margin*. That is, whether individuals correctly perceive the effects of the BE at the current incentive level r, as formalized in Proposition 2. The statistic $\bar{W}(r)$, however, captures individuals' perceptions of the difference in total costs, which includes, for example, differences in the fixed costs $K_i^1(0) - K_i^0(0)$. Proposition 1 provides a precise condition on $\frac{d}{dr}\bar{W}(r)$ under the null of correct perceptions, but it does not provide a precise condition on $\bar{W}(r)$. However, the impact on $\bar{W}(r)$ of additional treatments like opportunities for learning and feedback—as in our third experiment—can provide some insight about biases in the perception of total costs, and how those might differ from biases about the effects of the BE on the margin.

1.3 Uncertainty and Systematic Miscalibration

The test in Proposition 1 is a test of the assumption that individuals are neither systematically biased nor even uncertain about their attention cost functions. A plausible hypothesis is that individuals are not systematically biased, but are uncertain about the effects of BEs in new environments. To formalize this, let ξ denote an environmental parameter, such that the effects of the BE are given by $D(a = 1|r,\xi)$. Let each individual i receive a signal ζ_i about $D(a = 1|r,\xi)$, which is affiliated with ξ according to some joint distribution $H(\zeta,\xi)$, with marginals H_ζ and H_ξ. Suppose that individuals have an unbiased prior about the effects of the BE, centered around $\bar{D}(a = 1|r) := \int D(a = 1|r,\xi)dH_\xi(\xi)$, the average effect

across different environments. Then, if individuals are Bayesian, the martingale property of beliefs implies that individuals' posterior beliefs after receiving signal ζ_i, $\tilde{D}_i(a=1|r,\zeta_i)$, must also be unbiased when aggregating across environments ξ: $\mathbb{E}_{i,\xi}\left[\tilde{D}_i(a=1|r,\zeta_i)\right] = \bar{D}(a=1|r)$. By Proposition 2, this fact implies that $\mathbb{E}_{i,\xi}\left[\frac{d}{dr}\bar{W}(r|\xi) - \bar{D}(a=1|r,\xi)\right] = 0$. In particular, a key implication is that if individuals are Bayesians with unbiased priors, then although they might undervalue BEs in some situations ξ, they should not undervalue BEs *all the time*. In fact, if individuals undervalue BEs some of the time then they should *overvalue* them some of the time as well.

To make this idea more concrete, suppose that individuals have a normal prior with variance σ_0^2, and they receive a signal about $D(a=1|r,\xi)$ that is normally distributed around $D(a=1|r,\xi)$ with variance σ_1^2. Then individuals' perceptions, and thus by Proposition 2 their WTPs, satisfy

$$\begin{aligned}\frac{d}{dr}\bar{W}(r|\xi) &= \mathbb{E}_i\left[\tilde{D}_i(a=1|r,\zeta_i)|\xi\right] \\ &= (1-\theta)\bar{D}(a=1|r) + \theta D(a=1|r,\xi), \end{aligned} \tag{8}$$

where $\theta = \sigma_0^2/(\sigma_0^2+\sigma_1^2)$ is the degree of Bayesian shrinkage toward the prior mean. This implies that individuals undervalue the BE when it is more effective than average, in the sense that $D(a=1|r,\xi) > \bar{D}(a=1|r)$. But individuals also overvalue the BE when it is less effective than average, in the sense that $D(a=1|r,\xi) < \bar{D}(a=1|r)$. Thus, showing that individuals overestimate or underestimate the effects of a BE in one particular decision environment does not imply that the miscalibration is systematic.

In data sets where there is exogenous variation in conditions ξ that generate variation in the efficacy of the BE, equation (8) can be estimated simply through the linear regression model

$$\frac{d}{dr}\bar{W}(r|\xi) = \beta_0 + \beta_1 D(a=1|r,\xi). \tag{9}$$

The coefficient β_1 identifies θ, which implies that $\bar{D}(a=1|r) = \beta_0/(1-\beta_1)$.

More generally, equation (9) can also be used to explore systematically biased perceptions. For example, equation (9) can also be interpreted as a "meta-inattention" model in the spirit of the attribute-misperception model of Gabaix (2014), where $d = \beta_0/(1-\beta_1)$ is some default perception that people "anchor" on, but is not necessarily an unbiased representation of the average effects of the BE. To illustrate, suppose that it is known that the effects of reminders are generally non-negative.[8] Then detecting strongly positive effects

[8] I.e., the "rational" prior puts little weight on negative effects.

in some conditions suggests that $\bar{D}(a = 1|r) > 0$, and thus a finding that $\beta_0 = 0$ suggests that people systematically underestimate the effects of reminders because they anchor on the erroneous default perception of null effects.

2 Online Education Experiment

Our first experiment was designed around the Slutsky symmetry test in equation (3) of Proposition 1. It was run in the fall of 2018. We partnered with Code Avengers, an online platform for learning to code, to offer participants a free, eight-week course in three different programming languages (HTML/CSS, Javascript, and Web Dev).[9] Screenshots of all experimental instructions are in Screenshots Appendix E.1.

2.1 Design and Implementation

2.1.1 Participant Pool

We recruited students and recent alumni from six Philadelphia-area colleges using an email campaign. Enrollees were eligible to be included in our study if they reported in the on-boarding survey that they regularly used either Google Calendar or Apple's iCal as an electronic calendar. Perhaps due to the relative youth of the participant pool, usage rates were high, at around 60–70 percent. Recruitment resulted in a pool of 1373 study-eligible participants.[10]

2.1.2 Implementation

Just before the 8-week course began, participants received an introductory email with information on their treatment assignment. This email also contained a recommendation that participants aim to complete three, 15-minute sessions of the coding course per week, a prompt to encourage participants to make a plan for when they would do the coding lessons, and a link to make plans for working on the coding lessons, which would be created in their electronic calendars. Participants who were eligible for financial rewards were informed that they would be paid their cumulative earnings in the form of an Amazon gift card at the end of the 8-week period.

[9] These languages are commonly used tools for building modern web sites. See http://www.codeavengers.com for more details.

[10] Appendix Table B.1 presents characteristics of the participant pool. Females, first-years, and seniors were most likely to participate.

Over 90 percent of participants opened the initial emails informing them of the incentives they faced (i.e., their treatment), giving us confidence that most were aware of the incentives for which they were eligible. As expected from random assignment of treatment, email opening rates were very similar across treatments, ranging from 88 to 91 percent.

After the course had begun, all participants received a reminder email at the start of each week. The reminder email contained the same recommendation, planning prompt, and link to create plans as the initial email.

2.1.3 Experimental Design

The experiment consisted of a control group and five treatment arms, with varying levels of incentives for plan making and/or coding task completion. Participants assigned to the control group received the initial and reminder emails encouraging them to plan and complete the coding lessons and offering them the plan-making tool, but they were not eligible for financial rewards.

Those randomly assigned to the two *Pay-to-Plan* treatments received either $1 or $2 for making a plan for when to do their coding lessons that week (i.e., clicking the plan-making link within the weekly email). In the two *Pay-to-Code* treatments, participants received either $2 or $5 for completing three 15-minute sessions of the coding course during the week. Finally, participants in the *Combination* treatment arm were paid $1 for making a plan plus $2 if they completed three 15-minute sessions of the coding course during the week. Participants could earn these amounts each week, regardless of what they had done in previous weeks. In addition, making a plan did not restrict when a participant could do the coding lessons (i.e., participants in the *Pay-to-Code* and *Combination* treatments could complete the 15-minute sessions at any time during the week and still earn their coding-task incentives, regardless of whether or not they made a plan or when they had scheduled the three 15-minute sessions).

To measure plan making, we tracked whether a participant clicked on the provided plan-making link to create calendar events for when they planned to complete the 15-minute coding sessions.[11] Consistent with our theoretical framework, this observable plan-making is not the only available bandwidth enhancement (BE), or even the only available plan-making opportunity. For example, some participants might have other means of making plans or might directly edit their calendars without using our link. However, nearly 40

[11] When participants clicked on the plan-making link, they were given three default times, which they could change. This default ensured that as long as a participant clicked on the link, a calendar event would be created.

percent of the control group clicked to make a plan in the first week, despite receiving no financial rewards for doing so, and participants with higher incentives for completing the coding task were more likely to use the plan-making tool, implying that our plan-making tool was not a perfect substitute for the plan making individuals would do otherwise.[12] This may be because the act of making a plan by using our link generates an internal cue, as theorized in the implementation intentions literature (Gollwitzer and Sheeran, 2006).

To measure completion of the coding coursework, we received real-time, backend data from Code Avengers on the number of minutes participants spent actively working on their coding coursework each day. The session timer stopped running after approximately 30 seconds of inactivity within the course. Once they had completed 15 minutes of active work, participants were notified with a pop-up that congratulated them but did not prevent or discourage them from continuing.

2.2 Results

2.2.1 Empirical Framework

Our primary analysis focuses on measuring the effect of plan-making and coding-task incentives on plan making and coding task completion. We estimate treatment effects using regressions of the form:

$$y_{ict} = \beta T_{ict} + \alpha_c + \alpha_t + \gamma X_i + \varepsilon_{ict}, \tag{10}$$

where y_{ict} measures either plan making or completing at least $\tau \in \{0, 10, 20, 30, 40, 45, 50, 60\}$ minutes in week t for participant i at campus c. We include fixed effects α_c for campus interacted with student status (i.e., current student or alumni), which was the level at which we randomized. We also control for course week α_t and a vector of participant characteristics X_i, but random assignment implies that these additional controls should not affect our estimated treatment effects. Our preferred measure of treatment T_{ict} is value in dollars of the participant's incentive, which assumes a linear relationship between the incentive and behavior. We also consider a specification with indicators for different incentive sizes. We estimate regressions separately for the Pay-to-Plan sample, which includes the control

[12] Our theoretical framework only requires that the plan-making tool we offer is not a perfect substitute to other forms of planning individuals already undertake. Heterogeneity in attention cost functions accommodates the possibility that some participants who use our plan-making tool simply substituted from creating their own calendar reminders while others who use our plan-making tool would not have created a plan themselves.

group and the two *Pay-to-Plan* treatments, and the Pay-to-Code sample, which includes the control group and the two *Pay-to-Code* treatments.

2.2.2 Plan-making Incentives

In Table 1, we estimate the impacts of plan-making incentives on plan making and on coding task completion. In the context of our model, these specifications measure the derivatives of $\frac{d}{dp}Pr(a = 1|p,r)$ and $\frac{d}{dp}Pr(j = 1|p,r)$. The analysis sample includes 705 participants and eight pooled weekly observations per participant. In Panel A, we estimate the effect of plan-making incentives on the propensity to plan in week 1, weeks 1 to 4, and weeks 1 to 8. Multiple-week outcomes average the indicator for whether a participant made a plan (or completed the coding task) in each week. In Panel B, we estimate the effect of plan-making incentives on the propensity to complete at least 20 minutes or at least 45 minutes of coding during week 1, weeks 1 to 4, and weeks 1 to 8, respectively. Although our financial incentives were specifically for completing at least 45 minutes of the coding task (i.e., the three 15-minute sessions), we also include the 20-minute benchmark in the main tables and text to show robustness. Appendix Tables B.2 and B.3 consider other time thresholds: 0, 10, 30, 40, 50, and 60 minutes per week. Our interpretation of the results is consistent with the evidence from these alternative thresholds.

The results indicate strong impacts of plan-making incentives on plan making, and modest impacts of plan-making incentives on coding task completion. For each \$1 of plan-making incentive, participants increase their plan making by 11.6 percentage points ($se = 1.3$) on average over the eight weeks of the study, an increase of 140% relative to the control group mean of 8.2 percentage points. Plan-making effects are 18.0 percentage points ($se = 2.0$) in week 1, and 14.2 percentage points ($se = 1.4$) on average over weeks 1 to 4, which suggests an attenuated response over the course of the study. However, the control mean falls even more quickly, from 38.1% in week 1, to 15.0% in the first four weeks, to 8.2% over the full study, such that the relative impact of plan-making incentives increases over time. Panel A of Appendix Table B.4 shows the effects of the \$1 and \$2 plan-making incentives separately.

The treatment effect of plan-making incentives on coding task completion is more modest but still meaningful. Focusing on course completion of at least 45 minutes a week, we find that \$1 of plan-making incentive increases coding task completion by 3.8 percentage points ($se = 1.8$) in week 1, an increase of 22% relative to the control group mean of 17.4 percentage points. However, the effect declines to a marginally significant 1.7 percentage points ($se = 1.2$) over weeks 1 to 4, and to a statistically insignificant 0.6 percentage points

Table 1: The Effect of Plan-Making Incentives on Plan Making and Task Completion

A. The Effect on Plan Making (First Stage)

	(1) Week 1	(2) Weeks 1-4	(3) Weeks 1-8
Plan Incentive	0.180***	0.142***	0.116***
	(0.020)	(0.014)	(0.013)
Obs.	705	705	705
R^2	0.137	0.163	0.131
Control Mean	0.381	0.150	0.082
Controls	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes

B. The Effect on Coding Task Completion (Reduced Form)

	(1) >20 (1)	(2) >20 (1-4)	(3) >20 (1-8)	(4) >45 (1)	(5) >45 (1-4)	(6) >45 (1-8)
Plan Incentive	0.040**	0.028**	0.013	0.038**	0.017	0.006
	(0.020)	(0.013)	(0.011)	(0.018)	(0.012)	(0.009)
Obs.	705	705	705	705	705	705
R^2	0.057	0.049	0.051	0.036	0.035	0.041
Control Mean	0.280	0.212	0.158	0.174	0.156	0.116
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes	Yes	Yes	Yes

C. The Effect of Plan Making on Coding Task Completion (IV)

	(1) >20 (1)	(2) >20 (1-4)	(3) >20 (1-8)	(4) >45 (1)	(5) >45 (1-4)	(6) >45 (1-8)
Plan Making	0.221**	0.194**	0.114	0.213**	0.118	0.053
	(0.105)	(0.087)	(0.086)	(0.096)	(0.076)	(0.074)
Obs.	705	705	705	705	705	705
R^2	0.147	0.174	0.133	0.092	0.120	0.085
Control Mean	0.280	0.212	0.158	0.174	0.156	0.116
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes	Yes	Yes	Yes

This table estimates the effect of plan-making incentives ("Plan Incentive") on plan making and coding task completion. Panel A shows the effect of plan-making incentives (in dollars) on whether participants made a plan. Column (1) shows the effect of plan-making incentives in week 1 of the experiment. Column (2) shows the average effect for the weeks 1–4. Column (3) shows the average effect for all weeks. Panel B shows the effect of plan-making incentives on coding task completion. Columns (1-3) show the effect on an indicator variable for whether or not the participant worked on the task for more than 20 minutes: Column (1) estimates the effect over week 1, Column (2) over weeks 1–4, and Column (3) over all weeks. Columns (4-6) show analogous estimates for an indicator for whether or not the participant worked on the task for more than 45 minutes. Panel C shows the 2SLS estimates instrumenting for whether participants made a plan using the plan-making incentive as an instrument. The dependent variables are the same as in Panel B. Standard errors are in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

($se = 0.9$) over weeks 1 to 8. In Panel C, we combine the plan making and coding task completion estimates in an instrumental variables estimation of the effect of plan making on coding task completion. Making a plan increases the probability of coding task completion by 21 to 22 percentage points in week 1, an 81% to 124% increase relative to control group means. This large effect is precisely estimated for week 1 and weeks 1 to 4 but diminishes over the full experimental period. Overall, the results point to the value of plan making for people who have some intrinsic motivation to complete the coding sessions. Panel B of Appendix Table B.4 shows the effects of the $1 and $2 plan-making incentives separately.

The decrease in treatment effects over time is not surprising, as many participants appear to attrit out of the coding course. Appendix Figure B.1 plots control group means for plan-making and coding task completion over the weeks of the experiment. Engagement in the first two weeks of the study is relatively high in the absence of monetary incentives—control group participation hovers between 20 and 30 percent. However, many participants disengage from both the plan-making tool, which falls close to zero by week 3, and from continuing the coding course, which falls to 10 percent participation by week 5, suggesting that motivation for the coding course diminished over time. In the context of our model, this implies that participants' estimates of $B(r)$ diminished as participants received additional signals about the course.[13]

2.2.3 Coding-task Incentives

Table 2 estimates the impacts of coding-task incentives on plan making and coding task completion. In the context of our model, these specifications measure the derivatives of $\frac{d}{dr}Pr(a = 1|p, r)$ and $\frac{d}{dr}Pr(j = 1|p, r)$. The analysis sample includes 714 participants and eight pooled weekly observations per participant. Following the structure of Table 1, in

[13]Note that this by itself does not imply a deviation from optimal Bayesian decision making. As a simple illustration, suppose that for each participant, the beliefs about $B(r)$ take the form of a Bernoulli random variable that takes on the values $\bar{B} > 0$ with probability 0.2 and $\underline{B} < 0$ with probability 0.8, such that $0.2\bar{B} + 0.8\underline{B} > 0$. Then participants would initially sign up given the positive expectation of $B(r)$. But if the realizations of B are independently distributed across participants, 80 percent of them would attrit after discovering that $B = \underline{B}$. And if the realizations of B are positively correlated across participants, then in "bad" states the number of participants attriting could be much larger than what participants initially expected. For example, if the realizations are perfectly correlated, and the prior is that $Pr(B = \bar{B}) = 0.8$, then in the state $B = \underline{B}$ the participants who attrit will have ex-ante expected to complete the course with 80 percent chance. This does not pose a threat to our theoretical results about tests of optimal valuation of BEs because the experiment involves weekly measures of engagement with the planning tool. Thus, even if participants initially overestimated their enthusiasm about the course, they had the opportunity to adjust those expectations before the next planning decision.

Panel A we estimate the effect of coding-task incentives on the propensity to plan in week 1, weeks 1 to 4, and weeks 1 to 8. In Panel B, we estimate the effect of coding-task incentives on the propensity to complete at least 20 minutes or at least 45 minutes of coding during week 1, weeks 1 to 4, and weeks 1 to 8, respectively.

Coding-task incentives have substantial effects on coding task completion, as shown in

Table 2: The Effect of Coding-Task Incentives on Plan Making and Task Completion

A. The Effect on Plan Making

	(1) Week 1	(2) Weeks 1-4	(3) Weeks 1-8
Task Incentive	0.025***	0.010**	0.007**
	(0.009)	(0.004)	(0.003)
Obs.	714	714	714
R^2	0.050	0.058	0.049
Control Mean	0.381	0.150	0.082
Controls	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes

B. The Effect on Coding Task Completion

	(1) >20 (1)	(2) >20 (1-4)	(3) >20 (1-8)	(4) >45 (1)	(5) >45 (1-4)	(6) >45 (1-8)
Task Incentive	0.038***	0.031***	0.025***	0.035***	0.024***	0.020***
	(0.009)	(0.006)	(0.005)	(0.008)	(0.006)	(0.005)
Obs.	714	714	714	714	714	714
R^2	0.043	0.059	0.069	0.041	0.057	0.075
Control Mean	0.280	0.212	0.158	0.174	0.156	0.116
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes	Yes	Yes	Yes

This table estimates the effect of coding-task incentives ("Task Incentive") on plan making and coding task completion. Panel A shows estimates of the effect of coding-task incentives (in dollars) on whether or not participants made a plan. Column (1) shows the effect of coding-task incentives week 1 of the experiment. Column (2) shows the average effect over weeks 1–4. Column (3) shows the effect over all weeks. Panel B shows the effect of coding-task incentives on coding task completion. Columns (1-3) show the effect on an indicator variable for whether or not the participant worked on the task for more than 20 minutes: Column (1) estimates the effect over week 1, Column (2) over weeks 1–4, and Column (3) over all weeks. Columns (4-6) show analogous estimates for an indicator for whether or not the participant worked on the task for more than 45 minutes. Standard errors are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Panel B. We estimate that each \$1 of coding-task incentive increases completion rates for 45-minutes in week 1 by 3.5 percentage points ($se = 0.8$), an increase of 20% relative to the control group mean of 17.4 percentage points. For the \$2-incentive and \$5-incentive groups, this coefficient implies an increase in the probability of coding task completion of 7 and 17.5 percentage points, respectively, or 40% and 101% relative to the control mean of 17.4 percentage points. Again, the treatment effects diminish over time to 2.4 percentage points ($se = 0.6$) per \$1 over weeks 1 to 4, and to 2.0 percentage points ($se = 0.5$) per \$1 over the eight weeks of the study.[14]

A more novel result is that coding-task incentives also increase the probability of plan making, as shown in Panel A. Column 1 shows that for each \$1 of coding-task incentive, participants increase their plan making by 2.5 percentage points ($se = 0.9$) in week 1, by 1.0 percentage point ($se = 0.4$) in weeks 1 to 4, and by 0.7 percentage points ($se = 0.3$) over the eight weeks of the study. Relative to the control group means of 38, 15, and 8 percentage points, these correspond to plan making increases of 6.6%, 6.7%, and 8.5% per \$1 of plan-making incentive.

2.2.4 Symmetry Test

Participants clearly recognize the potential value of plan making in helping them achieve their coding course participation. But do they value plan making enough? To answer this question, we compare the cross-price elasticities estimated in the Pay-to-Plan and Pay-to-Code samples, implementing the test in Equation (3) of Proposition 1. The coefficients for \$1 of plan-making incentives on coding task completion are 0.039, 0.017, and 0.006 in week 1, weeks 1 to 4, and weeks 1 to 8, respectively. The analogous coefficients for \$1 of coding-task incentives on plan making are 0.025, 0.010, and 0.006. The difference in coefficients provides our first test of under-planning, delivering estimates of 0.014 ($se = 0.019$), 0.007 ($se = 0.012$), and -0.0004 ($se = 0.009$), respectively.[15] The positive sign of the differences, particularly in the early weeks of the study, hints at the possibility that

[14] We exclude the *Combination* treatment from our main analysis and separately evaluate whether this treatment exhibits complementarity effects (i.e., whether combining a \$1 plan-making incentive with a \$2 coding-task incentive induces plan making or coding effects that are significantly different from the \$1 *Pay-to-Plan* or \$2 *Pay-to-Code* treatments in isolation). For weeks 1 to 8, the *Combination* treatment effect on plan making is 26.7 percentage points ($se = 2.6$) compared to 23.9 percentage points ($se = 2.7$) for the \$1 *Pay-to-Plan* treatment (p-value of difference = 0.31). The *Combination* treatment effect on average course completion is 3.8 percentage points ($se = 2.4$) compared to 4.6 percentage points ($se = 2.1$) for the \$2 *Pay-to-Code* treatment (p-value of difference = 0.72). Thus, we find no statistically significant complementarity effect of the *Combination* treatment.

[15] Standard errors for coefficient differences are estimated via seemingly unrelated regression.

participants might undervalue plan making. However, the standard errors are too wide to draw strong conclusions from this data about whether participants plan optimally.

Figure 1 plots week-by-week coefficients for plan-making and coding-task incentives to illustrate how the effect of incentives evolves over the course of the experiment. The effect of coding-task incentives on plan making is consistently close to zero (after week 1) and tightly estimated. In contrast, the effect of plan-making incentives on coding task completion is positive for the first half of the study and then decays toward zero, with relatively wider confidence intervals.[16] This provides suggestive evidence of under-planning.

3 Online Survey-completion Experiment

Complementing our first experiment, we ran a survey-completion experiment on Amazon's Mechanical Turk platform (MTurk). The experiment is tightly tied to the test in equation (2) of Proposition 1, described in Section 1. The test states that for individuals who optimally invest in bandwidth enhancements (BEs), a $1 increase in the incentive for task completion must increase willingness to pay for such a technology by $1 times its efficacy (i.e., by the change in the probability of task completion due to the BE). Screenshots of all experimental instructions are in Screenshots Appendix E.2.

3.1 Design and Implementation

Participants were recruited to complete part 1 of the study each weekday between September 7 and September 24 of 2021. Our recruitment material informed potential participants that part 1 of the study would require 15 minutes of time immediately (for which participants were paid a guaranteed $2.50 and had the possibility of earning a bonus), and that they would be invited to participate in part 2 of the study at a later date for additional compensation by accessing a website provided to them in part 1 of the study.

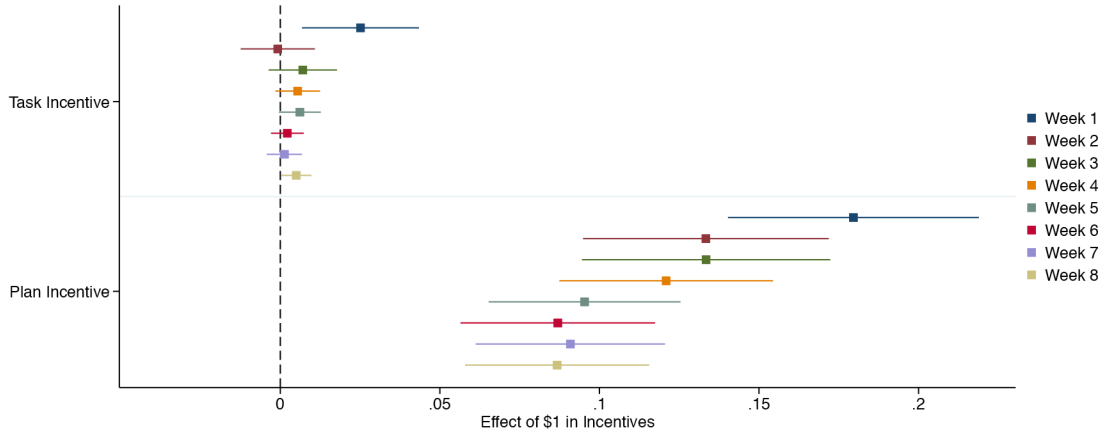
When participants clicked to begin the study, they were told that part 2 of the study—a survey that needed to be completed in one sitting of approximately 20 minutes—would only be available starting on some day in the future to be randomly determined during part 1. Participants were told that they would have a one-week window to complete it.

The first part of the study elicited participants' willingness to pay (WTP) for a set of three reminder emails (i.e., the BE in this experiment) that would come during the one-

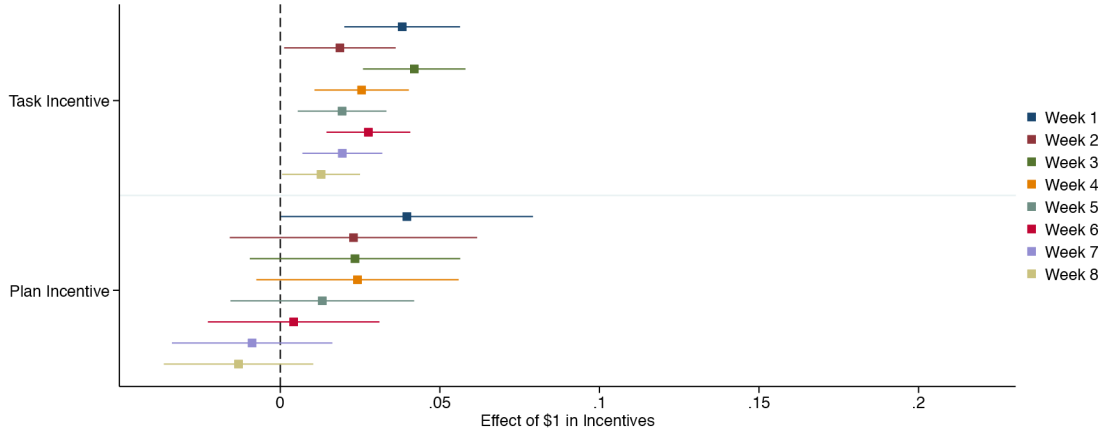
[16]The difference in standard errors across treatments is due to higher variance in coding-task incentives ($0, $2, and $5) relative to plan-making incentives ($0, $1, and $2).

Figure 1: The Effect of Incentives on Plan Making and Coding Task Completion

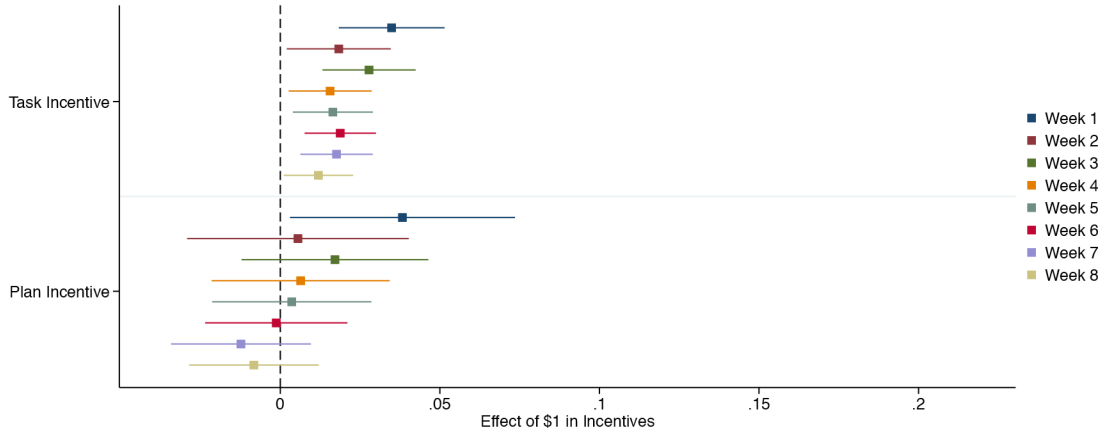
A. The Effect of Incentives on Plan Making



B. The Effect of Incentives on Coding Task Completion (>20 Minutes)



C. The Effect of Incentives on Coding Task Completion (>45 Minutes)



This figure shows estimates for the effect of incentives on plan making and coding task completion for each week of the study. Panel A shows estimates of the effect of incentives on whether or not participants made a plan. Panels B and C show the effect of incentives on completing at least 20 minutes and at least 45 minutes, respectively, of coding during the week. Whiskers report 95% confidence intervals around each estimate.

week window in which participants would be able to complete the survey. The goal was to generate data that would allow us to directly measure how much more participants were willing to pay for reminder emails as the incentive to complete the survey increased.

To ensure that participants understood the specific details of the reminder emails, we explained that the emails would come at 12 p.m. ET on the first, middle, and final days of the one-week window in which they could complete the survey.[17] Participants were told that emails would be sent using the MTurk email system—which MTurk uses for communicating with workers on its platform—so participants did not have to provide an email address and so the reminder emails would be unlikely to go to spam. Participants were told that the link to the survey would be included in the reminder emails so that initiating the survey would be as easy as clicking a link in the email. Participants were also explicitly told that they would not receive any reminders to complete the survey unless they were selected to receive these three reminder emails. We also clarified what the part-2 survey would look like (i.e., answering 40 hypothetical questions about gambles), and provided two example questions, in order to reduce ambiguity about the future tasks.[18]

Participants were informed that the survey would only be available starting in either 2 days, 1 week, 3 weeks, or 6 weeks, and that each delay was equally likely to be selected. Participants also learned that their incentive for completing the survey would be either \$3, \$4, \$11, or \$12. For ease of exposition, we refer to \$3 and \$4 as *low incentives* and \$11 and \$12 as *high incentives*. For each of the 16 combinations of the four possible incentive amounts and four possible delays, participants faced an incentivized multiple price list (MPL) that traded off part-1 bonus payments (up to \$4 for the low-incentive MPLs and up to \$12 for the high-incentive MPLs) against being sent the three reminder emails to complete the survey. Participants were informed that all possible bonus rewards, including part-1 and part-2 bonus payments, would be paid at the same point in time, after the one-week window to complete the survey ended.[19] Participants were randomly selected to

[17] For example, for participants who completed part 1 of the study on September 7, the 2-day-delay window was open from September 9–15. To any participants in the 2-day-delay group who were selected to receive reminder emails, we sent the emails at 12 p.m. ET on September 9, 12, and 15.

[18] In order to participate in the study, participants needed to correctly answer questions demonstrating their understanding of the compensation structure, the tasks in part 2 of the study, and the conditions for receiving reminders (i.e., they had to answer "True" to the statement: "You will not receive any reminders to complete part 2 of the study unless you are selected to get them in this part of the study."). Participants were also shown an MPL attention check screen that was used to remove participants who might click through the MPLs without reading the instructions.

[19] Part-1 bonus payments were paid out at the same time as any part-2 bonus payments, three days after the end of the one-week window to complete the survey, mitigating concerns that part-1 bonuses would be viewed as being paid immediately, which might have made them particularly valuable from the perspective of a quasi-hyperbolic discounter.

either first respond to the eight low-incentive MPLs (the "low-incentive block") or to the eight high-incentive MPLs (the "high-incentive block"). Within each block, the eight MPLs were shown to participants in a random order.

Because nuisance costs can lead participants to have negative WTP for the reminders, the MPL allowed participants to report both positive and negative willingness to pay for the reminder emails.[20] Participants were told that whichever incentive amount was randomly selected for them ($3, $4, $11, or $12) would be the bonus they would receive for completing the survey. In addition, they were told that for the randomly selected incentive amount, there was a 10% chance that one of the rows of that MPL would be randomly selected (each with equal probability) and that whatever the participant chose in that row would be implemented (i.e., they would receive whatever part 1 bonus payment was indicated in their choice, and they would receive the reminder emails if they chose the option on the left). Because testing Proposition 1 also requires estimating the effect of the reminder emails on completing the survey, we did not guarantee that one of the MPL rows would be selected. Instead, we randomized 45% of participants to receive the reminder emails and 45% of participants not to receive the reminder emails, regardless of their MPL choices. We use this random variation to estimate the effect of reminder emails on completing the survey. We randomly assign reminder emails in this way, and estimate the effect of reminders using this sample, in order to avoid potential selection bias that might arise if there were a correlation between WTP for reminders and the rate at which individuals completed the survey.

[20]Consistency on an MPL requires a participant to always choose the option on the left, always choose the option on the right, or switch from choosing the option on the left to choosing the option on the right in one row of the MPL. Our MPL was implemented to allow participants to choose a single cross-over point, thus enforcing consistency in choices. Use of single-cross-over MPLs is common in the experimental literature as they make the decision faster and easier for participants. The main concern is failing to identify participants who are clicking randomly through the study (i.e., those who would likely be identified as inconsistent on the MPL if required to make a selection in each row). This concern is mitigated in our setting because of our extensive attention checks.

3.2 Results

3.2.1 Sample

A total of 2743 individuals fully completed the first part of the study.[21] Additionally, we make the conservative sample restriction to limit all of our analysis to individuals whose WTPs were never top-coded at the smaller incentive (i.e., \$3 or \$11) or bottom-coded at the larger incentive (i.e., \$4 or \$12).[22] Mechanically, these top-coded and bottom-coded individuals cannot increase WTP when the task-completion incentive rises, which could lead to an attenuation bias in our estimates of how WTP for reminders changes with task-completion incentives. Given the wide range of values offered in the MPL, only 8.37% and 8.12% of responses were top-coded on the low-incentive and high-incentive MPLs, respectively, and 0.80% and 0.54% of responses were bottom-coded on the low-incentive and high-incentive MPLs, respectively. In what follows, we report on data from the remaining 2306 participants. Our restriction is conservative because it can only increase our estimates of how WTP for reminders varies with incentives; indeed, without this restriction, the point estimates are slightly lower.

3.2.2 Impact of Reminders on Survey Completion

As described in Section 3.1, we randomized 90% of participants to either get or not get the reminder emails, regardless of their reported WTP. This randomization allows us to generate an estimate of the effect of the reminders on survey completion at each delay. In addition, since we independently randomized the incentive level for completing the survey, we can estimate the effect of reminders at low and high incentive levels.

Figure 2 presents this data. Panel A shows the rate at which participants complete the survey at each delay, and by whether participants receive reminders. Panel B summarizes the treatment effect of receiving reminders at each delay and incentive level. Without reminders, completion rates decrease with delay at both high and low incentives. With re-

[21] This number does not include the 1854 participants who were automatically screened out of the study (and prevented from participating further) because they failed attention checks, ensuring our pool of participants understood the instructions in our experiment. It also excludes 36 individuals who were excluded for having an invalid MTurk ID or the 36 individuals who had technical issues in the display of MPL screens or recording of the data (e.g., being shown the wrong combination of incentives and delays or not receiving a link to the part-2 survey).

[22] We define *top-coded* participants as those who chose the option on the right in each row, indicating a WTP for reminders of more than \$4 (on the low-incentive MPLs) or \$12 (on the high-incentive MPLs). *Bottom-coded* participants chose the option on the left in each row, indicating a WTP for reminders of less than −\$4 or −\$12.

minders, however, the impact of delay on completion rates is much smaller. This translates into an increasing impact of reminders as delay increases, as shown in Panel B.

Table 3 quantifies these results. Column (1) shows that receiving the reminders increases the likelihood that participants complete the survey by 23 percentage points. The estimate on High Incentive shows that participants who receive high incentives to complete the survey are 7 percentage points more likely to complete the survey than those who receive low incentives. The coefficient on ln(P2 Delay) implies that participants are less likely to complete the survey as the delay increases. Column (2) shows that reminders have a significantly smaller effect at high incentives, but a significantly larger effect at longer delays. Column (3) shows that because reminders have very small effects in the short-delay / high-incentive conditions (Panel B of Figure 2), the impact of delay on the effects of reminders is particularly large in the high-incentive condition.[23] Columns (4) and (5) show that neither the point estimates nor the standard errors of the column (2) and (3) models change when (i) including fixed effects for when participants begin the study or (ii) doing two-way clustering by when participants started part 1 and by when part 2 was available.[24] Column (6) shows that a $1 change in the incentives is small enough to not significantly affect behavior, which allows us to utilize Corollary 1.

As formalized in Appendix A.5, the negative interaction between reminders and high incentives suggests that the effect of the high incentives on survey completion was at least in part due to individuals choosing a higher level of attention in the absence of reminders. In the language of our model, if incentives affected only the benefit $B(r)$, but not the likelihood of being attentive μ, then individuals would be more likely to complete the task conditional on being attentive. However, since reminders increase the likelihood of being attentive, there would then be a *positive* interaction between reminders and incentives. Instead, if higher incentives increase individuals' effort to be attentive even in the absence of reminders, then there is less need for reminders, leading to a negative interaction.

3.2.3 How WTP Changes with the Incentive to Complete the Survey

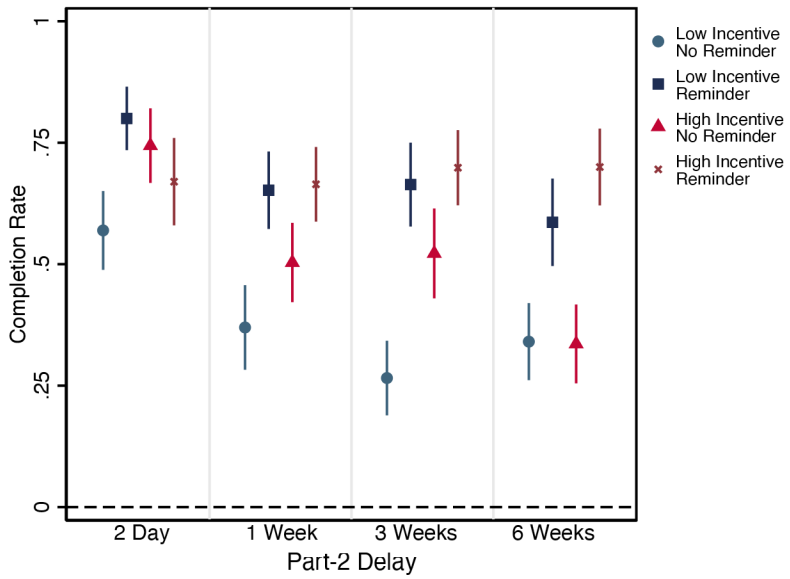
Figure 3 presents the average WTP for reminder emails for each part-2 incentive level at each of the four delays. Participants are willing to pay significantly more for reminders

[23] Appendix Table C.1 presents a less parametric regression analysis that separately estimates the effect of each of the possible delays, as well as its interaction with reminders, on survey completion.

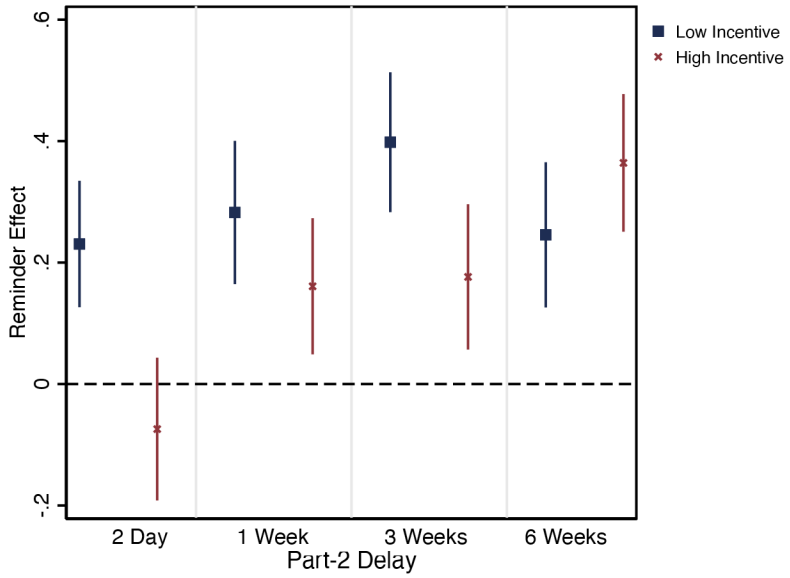
[24] We do not include fixed effects for part-2 survey availability because certain start dates are possible only in the 2-day condition, which means that the full set of fixed effects is not separately identified from the other covariates. However, because clustering by part-2 survey start date does not change standard errors, this is unlikely to matter.

Figure 2: Completion Rates and Treatment Effects of Reminders

A. Completion Rates by Incentive Level, Delay Type, and Reminders



B. Treatment Effect of Reminders by Incentive Level and Delay



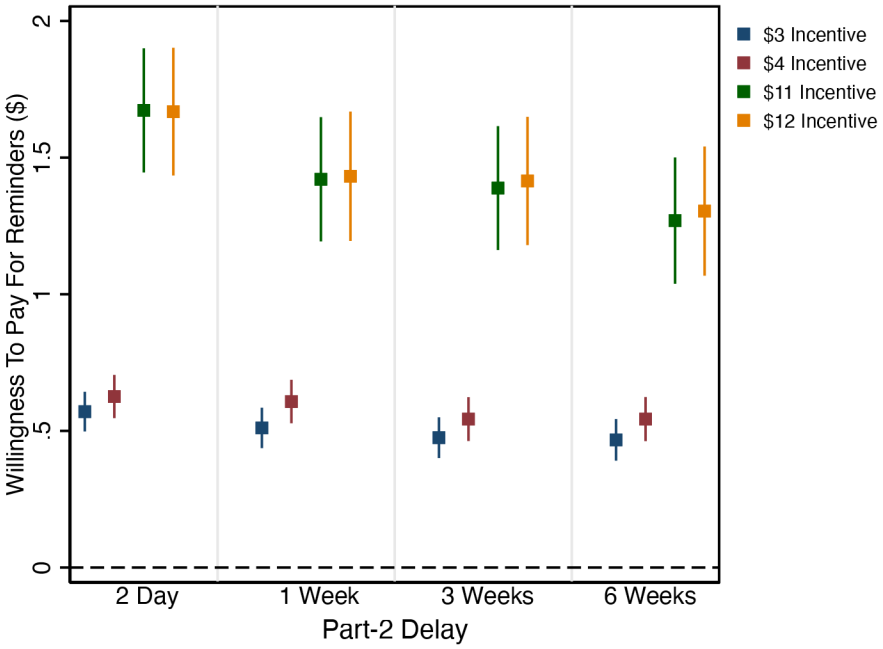
Panel A shows the part-2 survey completion rate and how it varies with the incentive for completing the survey, the amount of time after which part 2 of the survey became available (i.e., the part-2 delay), and whether the participant received reminders. Panel B shows the point estimates from a regression of the part-2 survey completion rate on whether the participant received a reminder. Both panels only include participants who were randomly assigned to receive or not receive reminders. The lines represent 95% confidence intervals.

Table 3: The Effect of Incentive, Delay, and Reminders on Part-2 Survey Completion

	Completed Part-2 Survey					
	(1)	(2)	(3)	(4)	(5)	(6)
Received Reminder	0.23***	0.12**	0.23***	0.12**	0.22***	0.23***
	(0.021)	(0.051)	(0.064)	(0.055)	(0.058)	(0.021)
High Incentive	0.07***	0.13***	0.21***	0.13***	0.20***	0.07***
	(0.021)	(0.029)	(0.066)	(0.031)	(0.063)	(0.021)
Ln(P2 Delay)	-0.07***	-0.10***	-0.08***	-0.10***	-0.08***	-0.07***
	(0.009)	(0.013)	(0.017)	(0.020)	(0.022)	(0.009)
Received Reminder × High Incentive		-0.13***	-0.38***	-0.11***	-0.32***	
		(0.042)	(0.094)	(0.039)	(0.070)	
Received Reminder × Ln(P2 Delay)		0.07***	0.02	0.07***	0.02	
		(0.018)	(0.025)	(0.020)	(0.025)	
High Incentive × Ln(P2 Delay)			-0.03		-0.03	
			(0.025)		(0.025)	
Received Reminder × High Incentive			0.11***		0.09***	
× Ln(P2 Delay)			(0.036)		(0.031)	
Extra $1						0.00
						(0.021)
Constant	0.57***	0.62***	0.59***	0.62***	0.59***	0.57***
	(0.028)	(0.036)	(0.045)	(0.054)	(0.057)	(0.029)
Observations	2,076	2,076	2,076	2,076	2,076	2,076
Number of Participants	2,076	2,076	2,076	2,076	2,076	2,076
S.E. Clustered by P1 & P2 Date				X	X	
P1 Date FE				X	X	

This table estimates how survey completion varies with reminders, the natural log of delay (in days), and whether participants are offered high incentives (i.e., $11 or $12) or low incentives (i.e., $3 or $4) to complete the survey. This table only includes participants who were randomly assigned to receive or not receive reminders. Columns (4) and (5) reproduce Columns (2) and (3) with fixed effects for the date that part 1 of the study was taken and with standard errors clustered for the date the participant completed part 1 and the date part 2 was made available to them. Column (6) reproduces Column (1) but estimates the impact of an extra $1 of incentive for completing the survey (i.e., the incentive being $4 or $12). Standard errors are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Figure 3: Willingness to Pay for Reminders



This figure shows mean willingness to pay for the reminders across different experimental conditions. The lines represent 95% confidence intervals with standard errors clustered at the participant level.

at the high incentives (i.e., $11 and $12, shown on the right of each panel) than at the low incentives (i.e., $3 and $4, shown on the left of each panel). Additionally, WTP for reminders is higher at the $4 incentive than at the $3 incentive.

We formalize the results from Figure 3 in Table 4, combining data from all four incentive levels and all four delays to estimate how average WTP changes with the incentive to complete the survey.[25] The coefficient on Extra $1 is the impact on willingness to pay for reminders of increasing the survey-completion incentive from $3 to $4. The coefficient on High Incentive compares average WTP at $3 to average WTP at $11. The coefficient on Extra $1 $\times$ High Incentive compares the impact on WTP of increasing the incentive by $1 when the incentives are high (i.e., going from $11 to $12) to the impact when incentives are low (i.e., going from $3 to $4).

Column (1) shows that as the incentive to complete the survey increases from $3 to $4, participants are on average willing to pay around 7 cents more for the reminders. However,

[25] Appendix Table C.2 replicates Table 4 for the 90% of participants who either receive or do not receive the reminder emails based on random assignment. As one would expect from the fact that this 90% is randomly selected, estimates are nearly identical to those in Table 4.

the interaction on *Extra* \$1 $\times$ *High Incentive* is negative and similarly sized, implying that when incentives are high, the extra dollar of incentive does not lead to an increase of WTP. That the coefficient on *High Incentive* is large and positive shows that WTP is on average much higher when incentives are high.[26]

Column (2) includes a covariate for the delay until the part-2 survey, Ln(P2 Delay), as well as the interaction Ln(P2 Delay) $\times$ Extra \$1. The coefficient on Ln(P2 Delay) reveals that participants are willing to pay less for reminders when the task is further out in the future, although the coefficient on Ln(P2 Delay) $\times$ Extra \$1 is directionally positive. Column (3) shows that these affects are amplified at the high incentive level. Column (4) returns to the specification in Column (1) but shows that controlling for the day participants completed part 1 of the study does not impact our estimates. Finally, Columns (5) and (6) replicate the specifications in (1) and (2) using a Tobit model to account for participants who were top-coded at the higher incentive levels within each group (\$4 or \$12). The results are quantitatively and qualitatively very similar in columns (5) and (6).[27]

3.2.4 Do Participants Invest in Reminders Optimally?

Taken together, the results show that, for a \$1 increase in the incentive for completing the survey, participants are on average willing to pay around 7 cents more for the reminders at low incentive levels but only 2 cents more at high incentive levels.

Part 1 of Proposition 1 states that if participants are optimally investing in the BE, then a small increase dr in incentives for completing the survey should increase WTP for the BE by approximately dr times the increase in the likelihood of survey completion due to the BE. As Column (6) of Table 3 shows, a \$1 change in incentives does not have a large effect on task completion, which implies that we can apply the test in part 1 of Corollary 1 to a \$1 change in incentives. On average, reminders had a 29 and a 16 percentage point effect on

[26] While not as natural a test of the theory since it spans a much larger increase in incentives, the 93 cent increase in WTP reflects an \$8 increase in the incentive level, or $\frac{93}{8} = 11.63$ cents per dollar, which is not that much larger than the 7 cent increase identified above. An additional difference that confounds this particular analysis, however, is that the MPL we use to elicit WTP for the high incentive levels was different (i.e., contained 33 rows where WTP increased in 75-cent increments) than the MPL for the low incentive levels (i.e., where WTP increased in 25-cent increments). For this reason as well, the estimate on *Extra* \$1 is the more natural test of the theory.

[27] Appendix Table C.3 shows that the results about WTP for reminders do not differ (at conventional levels of statistical significance) when we restrict to the first 2, 4, 6, or 8 MPL screens that participants encountered. This suggests that the within-subject design did not introduce demand or anchoring effects that altered our estimates. Appendix Table C.4 presents a less-parametric regression analysis that separately estimates the effect of each of the possible delays, as well as its interaction with the Extra \$1 covariate, on WTP for reminders.

Table 4: The Effect of Incentive and Delay on Willingness to Pay for Reminders

	WTP for Reminders ($)					
	(1)	(2)	(3)	(4)	(5)	(6)
Extra $1	0.07***	0.05	0.06*	0.07***	0.10***	0.07
	(0.017)	(0.051)	(0.036)	(0.017)	(0.018)	(0.052)
High Incentive	0.93***	0.93***	1.14***	0.93***	0.93***	0.93***
	(0.077)	(0.077)	(0.112)	(0.077)	(0.077)	(0.077)
Extra $1 × High Incentive	-0.06	-0.06	-0.08	-0.06	-0.07	-0.07
	(0.048)	(0.048)	(0.109)	(0.048)	(0.048)	(0.048)
Ln(P2 Delay)		-0.08***	-0.03***			-0.08***
		(0.022)	(0.012)			(0.022)
Extra $1 × Ln(P2 Delay)		0.01	0.00			0.01
		(0.021)	(0.014)			(0.021)
High Incentive × Ln(P2 Delay)			-0.09**			
			(0.035)			
Extra $1 × Ln(P2 Delay)			0.01			
× High Incentive			(0.041)			
Constant	0.51***	0.69***	0.59***	0.51***	0.50***	0.69***
	(0.032)	(0.057)	(0.041)	(0.032)	(0.032)	(0.057)
Observations	36,896	36,896	36,896	36,896	36,896	36,896
Number of Participants	2,306	2,306	2,306	2,306	2,306	2,306
Specification	OLS	OLS	OLS	OLS	Tobit	Tobit
P1 Date FE				X		

This table estimates the effect of incentive, the natural log of delay (in days), and having an incentive in the high-incentive pair on the willingness to pay for reminders. The extra $1 variable is an indicator for an incentive of $4 or $12. Participants in the high-incentive pair had a completion incentive of $11 or $12, compared to the low-incentive pair of $3 or $4. The "High Incentive" variable is an indicator for whether participants were in the high-incentive pair group. Column (1) shows OLS estimates for incentive, having an incentive in the high-incentive pair, and incentive interacted with having an incentive in the high-incentive pair; Column (2) maintains the specification in Column (1) and adds the natural log of delay; Column (3) adds an interaction between incentive, the natural log of delay, and an indicator for the high-incentive pair, as well as an interaction between the natural log of delay and the high-incentive pair, and an interaction between the natural log of delay and incentive; Column (4) shows Column (1) with fixed effects for the date that part-1 of the survey was taken; Column (5) uses the same variates as in Column (1) and shows Tobit estimates with censors at -$4 and $4 for the low-incentive group and censors at -$12 and $12 for the high-incentive group; Column (6) uses the same variates as in Column (2) and shows Tobit estimates with censors at -$4 and $4 for the low-incentive group and censors at -$12 and $12 for the high-incentive group. Standard errors, clustered at the participant level, are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

survey completion in the low-incentive and high-incentive conditions, respectively. This is significantly larger than the respective \$0.07 and \$0.02 changes in WTP with respect to an additional \$1 incentive in these conditions (Wald test $p < 0.01$).

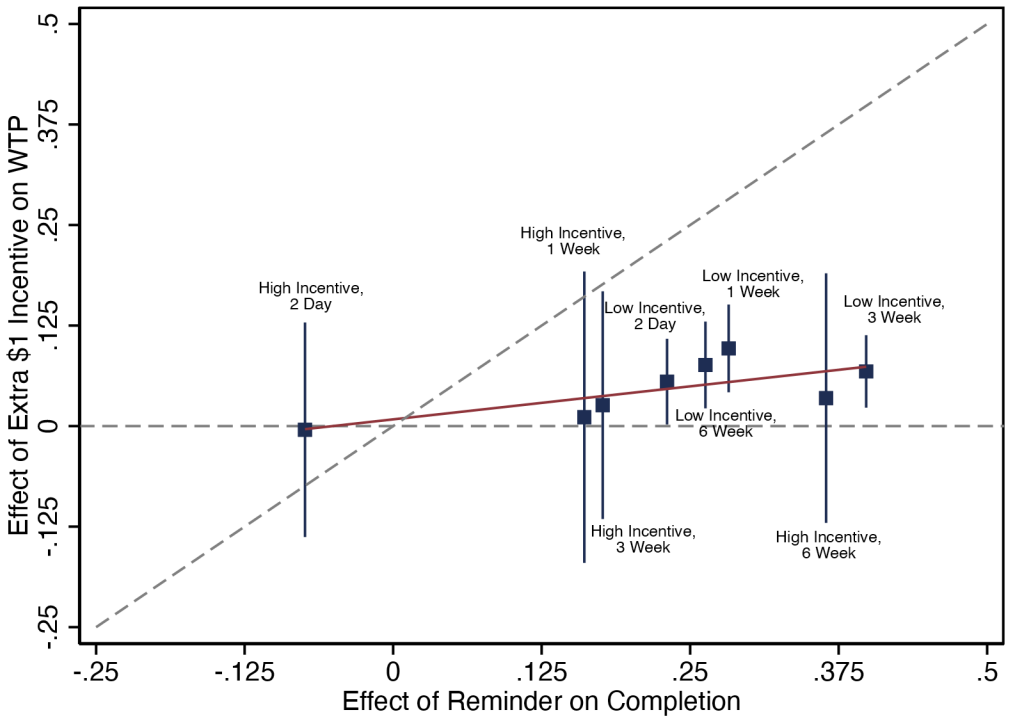
We can also apply Corollary 1 across each of the eight different incentive-delay pairs (i.e., $\{low \text{ and } high\} \times \{2 \text{ days, } 1 \text{ week, } 3 \text{ weeks, and } 6 \text{ weeks}\}$) generated in our experiment to estimate equation (9) from Section 1.3. A trivial application of the corollary is that under those same assumptions, equation (9) can be rewritten as

$$\bar{W}(r+\Delta|\xi) - \bar{W}(r|\xi) = \beta_0 + \beta_1 \frac{D(a=1|r,\xi) + D(a=1|r+\Delta,\xi)}{2}. \tag{11}$$

We estimate this equation by treating the eight different experimental conditions as variation in ξ. We then regress the change in WTP with respect to a \$1 change in incentives on the estimated effect of reminders, $D(a=1|r,\xi)$, in each condition. Formally, this procedure is equivalent to a two-stage least squares (2SLS) estimator, where the eight different experimental conditions are instruments for the effects of reminders, and where the dependent variable is the change in WTP with respect to a \$1 change in incentives.

Figure 4 provides a visualization of the second stage of this 2SLS estimator. On the x-axis, this figure shows the estimated effect of reminders on survey completion rates as reported in Panel B of Figure 2. On the y-axis, this figure shows the estimated increases in willingness to pay for the reminders as the incentive increases by \$1, together with the 95% confidence intervals of the estimates. If participants were optimally valuing the reminder technology, the WTP for the reminder would be on the 45-degree line (e.g., such that when reminders increase survey completion by 25 percentage points, the willingness to pay for the reminders increases by \$0.25 with a \$1 increase in incentives). Instead, our estimates are far from the 45-degree line. Estimated WTP is below the 45-degree line for seven of the eight estimates, and the 95% confidence intervals exclude the 45-degree line in six of the estimates. An estimate of equation (11) yields $\beta_0 = 0.01$ and $\beta_1 = 0.16$. That is, perceptions of the effects of reminders are attenuated toward a prior mean of approximately 0, by a factor of 84 percent. In line with the discussion in Section 1.3, this evidence may be more consistent with a model in which people's prior (or "default," in the sense of Gabaix, 2014) perceptions of the effects of the BE in this setting are systematically biased. For a prior mean of approximately 0 to be an unbiased prior, it would have to be that reminders often have no effect, and sometimes even have negative effects on task completion.

Figure 4: Effect of Reminders on Completion vs. Effect of Extra $1 on WTP



This figure displays estimates and standard errors for the effect of reminders on whether part 2 of the survey was completed, and the effect of an increase of $1 in the part-2 survey incentive on the willingness to pay for reminders. The figure only includes participants who were randomly assigned to receiving or not receiving reminders. The lines represent 95% confidence intervals, which are computed from the standard errors clustered by the date part 1 of the study was taken.

4 Learning Experiment

Results from our first two experiments reveal that individuals appear to undervalue bandwidth enhancements (BEs). One additional question is whether individuals can learn to value BEs through experience with them. We examine this in our third experiment. In addition, we show the versatility of our methods by using a design that builds on the Ambuehl et al. (2020) paradigm, which resembles other psychometric tasks that have been used to generate state-dependent stochastic choice data for testing rational inattention theories (Dean and Neligh, 2018; Caplin et al., 2020; Caplin, 2021). Screenshots of all experimental instructions are in Screenshots Appendix E.3.

4.1 Design and Implementation

We ran the learning experiment on Prolific Academic in October and November of 2021.[28] The study involved a series of tasks. In each task, participants were presented with an image that showed a series of equations that were either correct (e.g., $10 + 12 = 22$) or incorrect (e.g., $10 + 12 = 23$). Participants were asked to indicate whether the image contained more correct or more incorrect equations with an incentive for accuracy. The *baseline task* included an image with 100 equations about which participants were told that either 60% of the equations were correct (and 40% incorrect) or 40% of the equations were correct (and 60% incorrect).[29] Participants completed two blocks of seven tasks each, and at least three tasks in each block were baseline tasks. One task of the 14 tasks the participant completed was randomly selected, and participants were paid based on their accuracy in that task as described below.

Before completing each block of seven tasks, participants were asked for their willingness to pay to make the task easier (i.e., to take advantage of a BE). In the *length* arm, participants could shorten the task so that there were only 10 equations in the image, rather

[28] In the recruitment materials, potential participants were informed that the study would require 20 minutes of their time, for which they would receive a guaranteed $2.50. They were also informed that they would have the possibility of completing a bonus, and that the study had to be completed on a desktop or laptop computer using Chrome or Firefox as their web browser, which was necessary to ensure participants could see the tasks that were part of this study.

[29] The images were automatically constructed with randomly generated equations following the protocol in Ambuehl et al. (2020). The computer selected the number of equations (e.g., 100 in the baseline tasks) and randomly selected one of the possible fractions to be the percentage of correct equations (e.g., either 60% or 40% in the baseline tasks). The two numbers on the left side of the equation were each randomly selected from the range 1 to 99. For the correct equations, the true result appeared on the right side. For the incorrect equations, a number was randomly drawn from the range 1 to 5, which was either added to or subtracted from the true result at random. The equations were then shuffled for display.

than 100 equations. In the *discernibility* arm, participants could make the fraction of correct equations either 95% correct (and 5% incorrect) or 5% correct (and 95% incorrect), rather than 60% or 40%. Of the seven tasks in the block, three were baseline tasks and three were easy tasks (i.e., shorter tasks in the *length* arm and more discernible tasks in the *discernibility* arm). If the remaining task in the block was randomly selected for payment, then the participant's choices (i.e., responses on a multiple price list, as described below) would determine whether that task was a baseline task an easier task.[30]

Following the design of our survey-completion experiment, we elicited willingness to pay to reduce attention costs for different incentives for accuracy in the task. Participants were informed that the computer would randomly and independently select an accuracy bonus—paid to a participant for providing an accurate answer in the task—of either \$2, \$3, or \$4 for each block of seven tasks.[31] This procedure allowed us to elicit a participant's willingness to pay to make the task easier for three different accuracy bonus levels (i.e., before they knew which accuracy bonus would be relevant for that block). In particular, we elicited participants' willingness to pay to make the task easier using a set of three multiple price lists (MPLs). The interface of the MPLs was similar to that in the survey experiment.

A key design feature of this experiment is the opportunity to learn about the value of making the task easier. We do this by having participants experience both baseline and easy tasks in the first block of seven tasks (the baseline and easy tasks were presented in a random order). Participants then repeat the exercise—providing willingness to pay to make the task easier at each of the three accuracy incentive levels—for the second block of tasks.

The amount of feedback participants receive about their performance in the first block varies by treatment. In the *control* treatment, participants were not provided with information about their performance in the first block. In the *feedback* treatment, by contrast, participants were told the fraction of baseline and easy tasks they answered accurately. Moreover, participants were told how those accuracy levels translated to expected earnings at the three different levels of incentives.

In all MPLs, the options for willingness to pay ranged from $-\$4$ to \$4. This range is analogous to the range in the survey-completion experiment, where the highest possible MPL amount corresponded to the highest reward for task completion. Analogous to the

[30] If the task was not randomly selected for payment, then the difficulty of the task would be chosen at random. This protocol ensured that the multiple price list responses only affected the task and generated bonus payments in the case when participants were going to be paid based on their accuracy on that particular task.

[31] At the start of the block, participants were told which accuracy bonus had been randomly selected to apply for the block of seven tasks. Each task screen also reminded participants of the accuracy bonus. This bonus was paid if the participant provided an accurate answer on the task randomly selected for payment.

survey-completion experiment, we make the same conservative sample restriction to limit to individuals who were never top-coded at the smallest incentive (i.e., \$2) or bottom-coded at the largest incentive (i.e., \$4). The logic behind this restriction is the same as in the previous experiment: since these individuals cannot increase their willingness to pay as the accuracy bonus increases, it is possible that including these participants could lead to an attenuation bias in estimates of the effect of the accuracy bonus on willingness to pay for a BE.

4.2 Results

Our analysis involves the 1465 participants who completed the study, and who were not top- or bottom-coded as described above.[32]

4.2.1 Impact of Task Difficulty on Accuracy

To analyze the causal effect of task difficulty on performance, we exclude data from tasks that were potentially endogenously determined by participants' WTP. That is, we exclude cases in which the "remaining task" in the block was selected to be the task that counts, and thus was affected by participants' preferences.

Table 5 presents OLS regressions that quantify the impacts of task difficulty on the likelihood of correctly identifying whether there are more correct or incorrect equations in the picture. The first two columns analyze performance in the first block, while the latter two columns analyze performance in the second block. Columns (1) and (3) show that performance did not differ significantly between the two blocks. In both blocks, participants correctly completed the baseline task approximately 70 percent of the time, were approximately 26 percentage points more likely to complete it correctly when the task was more discernible (i.e., had either 95% or 5% correct calculations), and were approximately 19 percentage points more likely to complete it correctly when the task was shorter (i.e., had only 10 equations). On average, the discernibility effect was moderately larger than the length effect: by 6.5 percentage points in block 1 (Chi-square test $p = 0.00$) and by 4.9 percentage points in block 2 (Chi-square test $p = 0.00$). Columns (2) and (4) show that variation in our incentives did not have a significant effect on performance. This is consistent with the survey-completion experiment, where a \$1 change in incentives was too small

[32]This number does not include the 125 participants who were automatically screened out of the study (and prevented from participating further) because they failed attention checks, ensuring our pool of participants understood the instructions in our experiment. This number also does not include 13 individuals who encountered technical glitches.

Table 5: The Effect of Length and Discernibility on Getting a Task Correct

	Answered Task Correctly			
	(1)	(2)	(3)	(4)
More Discernible	0.26***	0.25***	0.26***	0.27***
	(0.010)	(0.008)	(0.010)	(0.008)
Shorter	0.19***	0.20***	0.21***	0.19***
	(0.011)	(0.009)	(0.012)	(0.010)
Length Arm	0.01		-0.04**	
	(0.015)		(0.015)	
Incentive ($)		-0.00		0.01
		(0.006)		(0.006)
Constant	0.71***	0.72***	0.71***	0.67***
	(0.010)	(0.018)	(0.010)	(0.018)
Observations	10,143	10,143	10,157	10,157
Number of Participants	1,465	1,465	1,465	1,465
Block	1	1	2	2

This table estimates the effect of shorter length (i.e., 10 calculations) and increased discernibility (i.e., 95% or 5% correct) on getting a task correct in block 1 and block 2. Tasks that had their difficulty determined by a participant's MPL choices have been excluded. The columns correspond to different regression specifications and blocks: Column (1) shows OLS estimates in block 1, Column (2) shows OLS estimates including the incentive level in dollars in block 1, and Columns (3) and (4) show analogous specifications for block 2. Standard errors, clustered at the participant level, are shown in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

to have a significant effect, despite behavior being overall elastic (and thus responsive to large changes). This result implies that the incentive changes we analyze are sufficiently small to utilize Corollary 1.

Data on decision times is consistent with the baseline task being more difficult. On average, participants spent 87 seconds, 44 seconds, and 37 seconds on the baseline, shorter, and more discernible tasks, respectively. Appendix Figure D.2 presents the CDFs of response times across the three types of tasks. Appendix Figure D.1 shows that participants who spent more time on the tasks were more likely to answer them accurately, particularly in the more difficult, baseline task.

4.2.2 Willingness to Pay to Simplify Tasks

Table 6 presents OLS regressions that estimate how participants' WTP to make the tasks easier varies with accuracy incentives and other experimental manipulations. Column (1) shows that participants in the *length* arm increased their WTP to make the task shorter by approximately \$0.10 for every dollar of extra accuracy incentive in block 1. However, this effect goes down to approximately \$0.03 in block 2, which is a significant difference of approximately \$0.07 (Chi-square $p = 0.06$). In the *discernibility* arm, each \$1 of accuracy incentive increases WTP by \$$-0.01$ and \$0.03 cents in blocks 1 and 2, respectively. Neither of these is statistically significantly different from zero, nor are they different from each other. In block 1, the difference between the coefficients on incentive in the length and discernibility arms is \$0.11, which is statistically significant ($se = 0.045$, Chi-square $p = 0.017$).

Given that simplifying the task increased accuracy by approximately 19 to 26 percentage points in both arms, Proposition 1 implies that WTP for simplifying the task should increase by approximately \$0.19 to \$0.26. This is higher than the effects reported in Table 6. In each of the four block-arm pairs, the WTP increase is significantly smaller than the theoretical benchmark (Wald test $p < 0.01$ in all arms).

Columns (1) and (2) thus reveal three key insights. First, people in this experiment undervalue BEs, as in our first two experiments. Second, participants initially undervalue discernibility improvements more than length improvements. An ex-post rationalization is that decreasing the length of a task is a simple and relatively common form of simplification that most people are familiar with, whereas increasing discernibility is a more abstract and less-common form of simplification. Third, experience does not bring people's decisions more in line with the correct-perceptions benchmark formalized in Proposition 1. Directionally, people value discernibility improvements more in block 2, but this is not statistically significant at conventional levels. In fact, there is stronger evidence that in block 2 of the length arm, there is *more* deviation from the Proposition 1 benchmark than in block 1. The difference between the coefficients on Incentive $\times$ Block 2 in columns (1) and (2) is -0.11 ($se = 0.052$, Chi-square $p = 0.04$).

Columns (3) and (4) provide insight into why the deviation increases with experience in the length arm. As shown in column (3), the deviation increases primarily among the participants who received the feedback treatment. For participants in the control treatment, the coefficient on incentive increases by an insignificant 0.01 ($se = 0.050$). However, as the coefficient on the interaction Incentive $\times$ Block 2 $\times$ Feedback shows, the impact of

Table 6: The Effect of Incentive, Block, and Feedback on Willingness to Pay for Easier Tasks

	Willingness to Pay ($)					
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive ($)	0.10***	-0.01	0.04	0.02	0.07	0.02
	(0.034)	(0.030)	(0.045)	(0.042)	(0.049)	(0.072)
Incentive ($) × Block 2	-0.07*	0.04	0.01	0.03	0.00	0.01
	(0.037)	(0.037)	(0.050)	(0.052)	(0.063)	(0.074)
Incentive ($)			0.11	-0.07	0.12	0.10
× Feedback			(0.067)	(0.060)	(0.089)	(0.099)
Incentive ($) × Block 2			-0.16**	0.01	-0.22**	-0.11
× Feedback			(0.074)	(0.074)	(0.100)	(0.108)
Block 2	0.15	-0.18	-0.09	-0.20	-0.16	-0.04
	(0.125)	(0.123)	(0.177)	(0.176)	(0.235)	(0.259)
Block 2 × Feedback			0.49*	0.05	0.87**	0.17
			(0.249)	(0.247)	(0.341)	(0.355)
Feedback			-0.22	0.22	-0.36	-0.11
			(0.248)	(0.213)	(0.342)	(0.353)
Constant	0.20	0.53***	0.31*	0.42***	0.32	0.31
	(0.124)	(0.107)	(0.172)	(0.147)	(0.212)	(0.260)
Observations	3,996	4,794	3,996	4,794	1,788	2,208
Number of Participants	666	799	666	799	298	368
Block 1 Acc. Diff.	All	All	All	All	≤ 0	> 0
Arm	Length	Discernibility	Length	Discernibility	Length	Length

This table estimates the effect of accuracy incentives, block order, and whether the participant received performance feedback on willingness to pay for an easier task (i.e., a shorter task in the length arm and a more discernible task in the discernibility arm). Columns (5) and (6) restrict participants by their block-1 accuracy difference between the baseline and easy tasks, which equals the difference between the percentage of easy tasks and baseline tasks answered correctly in block 1. The mean block-1 accuracy differences for participants in Columns (5) and (6) are -0.06 and 0.40, respectively. Column (1) shows OLS estimates for incentive, block, and the interaction of incentive and block order for participants in the length arm; Column (2) repeats this analysis in the discernibility arm; Column (3) maintains the specification in Column (1) and the restriction to participants in the length arm while adding whether feedback was received and the interactions between feedback, block, and incentive; Column (4) shows the OLS estimates in Column (3) for participants in the discernibility arm; Column (5) shows the OLS estimates in Column (3) for participants in the length arm who had a block-1 accuracy difference of less than or equal to 0; Column (6) shows the OLS estimates in Column (3) for participants in the length arm who had a block-1 accuracy difference of greater than 0. Standard errors, clustered at the participant level, are shown in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

experience is a statistically significant -0.16 ($se = 0.074$, $p = 0.034$) for participants in the feedback treatment. By contrast, feedback has no effect on participants in the discernibility arm, suggesting that participants in that arm have a very strongly held prior that discernibility would not affect their performance.

Columns (5) and (6) further explore the negative effect of feedback in the length arm. Column (5) restricts to participants who did not perform better on the shorter tasks than on the baseline tasks in block 1; column (6) restricts to participants who did perform better. The coefficient on Incentive $\times$ Block 2 is nearly identical in those two columns, implying that these two groups of participants were not differentially affected by experience in the control condition. However, the negative coefficient on Incentive $\times$ Block 2 $\times$ Feedback is twice as large in magnitude in column (5) than in column (6). This result suggests that a reason for the negative effect of the feedback treatment is participants being disappointed to learn that their block-1 performance was not higher on shorter tasks.[33]

The results about WTP in columns (3)–(6) of Table 6 cannot be explained by differences in block-2 performance among the different groups analyzed in those columns. Columns (1) and (2) of Appendix Table D.3 show that neither experience nor feedback altered participants' performance on the baseline versus easy tasks. Columns (3) through (5) of Appendix Table D.3 restrict to the length arm, and show that: (i) consistent with mean reversion, participants who did not perform better on the shorter tasks than on the baseline tasks in block 1 *improved* their relative performance on the easy tasks in block 2, and (ii) feedback did not reduce the difference in block-2 performance between the easy and baseline tasks among these participants. The overall negative effect of the feedback treatment suggests participants updated their beliefs in a quasi-Bayesian manner by overweighting disappointing experiences.

Interestingly, columns (3) and (5) of Table 6 also suggest that, while experience and feedback led participants to underestimate the effect of task simplification on their performance at the incentives in the experiment, it did increase their overall WTP to simplify the task. This result illustrates the theoretical discussion in Section 1.2.1 about how accurate perceptions of total costs (including fixed costs) are not characterized in Proposition 1—perceptions of the effects of the BE at the current incentive level r is not a sufficient statistic for perceptions of total costs. Appendix Table D.4 shows that participants who did not perform better on the shorter tasks than on the baseline tasks in block 1 spent almost

[33] Alternatively, it could be consistent with the suggestive evidence from column (6) of Appendix Table D.3 that feedback may have decreased performance for these participants—which suggests that these participants overall chose to adopt attention strategies that would decrease relative performance in the shorter task.

two minutes longer on the baseline versus easy tasks in block 1. Subsequently, they spent approximately 30 to 45 seconds less time on the baseline versus easy tasks in block 2, relative to the participants who did perform better on the shorter tasks in block 1. This suggests that the participants who did not perform better on the shorter tasks in block 1 also incurred significantly larger total costs on the baseline tasks. The feedback treatment may have helped prime this realization by inducing participants to further reflect on the differences between the baseline and shorter tasks.[34] This illustrates the Section 1.2.1 discussion about how additional treatments such as our feedback treatment can provide additional insights into perceptions about total costs, complementing the tests in Proposition 1.[35]

5 Conclusion

While a large and growing literature shows that attention-increasing interventions such as reminders and plan-making tools can have significant effects on economically important behaviors, this literature rarely asks the question of whether individuals value and deploy these tools optimally. This paper addresses this question with three theory-driven, quantitative tests. We find that individuals' demand for attention-increasing tools is *qualitatively* consistent with the predictions of rational inattention but is quantitatively inconsistent with the null of correct perceptions and optimal use of technologies to reduce attention costs. This suggests that individuals are uncertain and/or systematically biased about their attention cost functions. While this under-valuation of bandwidth enhancements may be context dependent, our methods can be applied more broadly to explore how individuals value attention-increasing technologies across various domains.

Our methods are immediately portable to other settings where the impact of reminders and planning prompts has already been documented, such as in medical compliance, savings, loan repayment, and voting (see footnote 3 for references). In addition, as exemplified by our third experiment, our methodology can be used to test whether people understand their production functions for attention-consuming tasks in field settings such as those of Dean (2019), Kaur et al. (2021), or Bessone et al. (forthcoming). More generally, our tests could be applied to any setting that involves domains of behavior that feature "intermedi-

[34] Appendix Table D.4 also shows that feedback treatment had a small negative effect on the time taken on the baseline tasks in block 2 for participants who performed better on the shorter tasks than on the baseline tasks in block 1. There is no effect on participants who did not perform better on the shorter tasks in block 1.

[35] Appendix Tables D.1 and D.2 replicate Table 6 using Tobit models and dropping participants with the 10 percent fastest average task times by the length and discernibility arms separately. The results are quantitatively and qualitatively similar.

ate" actions. For example, our methods could be used to quantify whether students fully understand the relationship between studying and test performance, whether individuals understand the link between education and earnings, or whether individuals properly invest in "good habits."

References

Abaluck, Jason and Abi Adams-Prassl, "What Do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses," *The Quarterly Journal of Economics*, 2021, *136* (3), 1611–1663.

Altmann, Steffen and Christian Traxler, "Nudges at the Dentist," *European Economic Review*, 2014, *72*, 19–38.

__ **, Andreas Grunewald, and Jonas Radbruch**, "Interventions and Cognitive Spillovers," *Review of Economic Studies*, forthcoming.

Ambuehl, Sandro, Axel Ockenfels, and Colin Stewart, "Who Opts In?," *Working Paper no. 7091, CESifo*, 2020.

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka, "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition," *American Economic Review*, June 2016, *106* (6), 1437–75.

Bernheim, B. Douglas and Dmitry Taubinsky, "Behavioral Public Economics," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics*, Vol. 1, New York: Elsevier, 2018, pp. 381–516.

Bessone, Pedro, Gautam Rao, Frank Schilbach, Heather Scofield, and Mattie Toma, "The Economic Consequences of Increasing Sleep among the Urban Poor," *The Quarterly Journal of Economics*, forthcoming.

Bronchetti, Erin Todd, David B. Huffman, and Ellen Magenheim, "Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination," *Journal of Economic Behavior and Organization*, 2015, *116*, 270–291.

Calzolari, Giacomo and Mattia Nardotto, "Effective Reminders," *Management Science*, 2017, *63* (9), 2915–2932.

Caplin, Andrew, "Measuring and Modeling Attention," *Annual Review of Economics*, 2016, *8*, 379–403.

__ , "Economic Data Engineering," *Working Paper no. 29378, National Bureau of Economic Research*, 2021.

__ **and Mark Dean**, "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, July 2015, *105* (7), 2183–2203.

__ **, Daniel Csaba, John Leahy, and Oded Nov**, "Rational Inattention, Competitive Supply, and Psychometrics," *Quarterly Journal of Economics*, 2020, *135* (3), 1681–1724.

__ **, Mark Dean, and John Leahy**, "Rational Inattention, Optimal Consideration Sets, and Stochastic Choice," *Review of Economic Studies*, 2019, *86* (3), 1061–1094.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky, "The Limits of Simple Implementation Intentions: Evidence from a Field Experiment on Making Plans to Exercise," *Journal of Health Economics*, 2018, *62*, 95–104.

Carvalho, Leandro and Dan Silverman, "Complexity and Sophistication," *Working Paper no. 26036, National Bureau of Economic Research*, 2019.

Castleman, Benjamin L. and Lindsay C. Page, "Freshman Year Financial Aid Nudges: An Experiment to Increase FAFSA Renewal and College Persistence," *Journal of Human Resources*, 2016, *51* (2), 389–415.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

Damgaard, Mette Trier and Christina Gravert, "The Hidden Costs of Nudging: Experimental Evidence from Reminders in Fundraising," *Journal of Public Economics*, 2018, *157*, 15–26.

Dean, Joshua T., "Noise, Cognitive Function and Worker Productivity," *Working Paper*, 2019.

Dean, Mark and Nathaniel Neligh, "Experimental Tests of Rational Inattention," *Working Paper*, 2018.

DellaVigna, Stefano and Devin Pope, "What Motivates Effort? Evidence and Expert Forecasts," *The Review of Economic Studies*, 2017, *85* (2), 1029–1069.

Ericson, Keith, "Forgetting We Forget: Overconfidence and Memory," *Journal of the European Economic Association*, 2011, *9* (1), 43–60.

_ , "On the Interaction of Memory and Procsastination: Implications for Reminders, Deadlines, and Empirical Estimation," *Journal of the European Economic Assocation*, 2017, *15* (3), 692–719.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

_ , "Behavioral Inattention," in Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics*, Vol. 2, Elsevier, 2019, pp. 261–343.

_ , David Laibson, Guillermo Moloche, and Stephen Weinberg, "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," *American Economic Review*, 2006, *96* (4), 1043–1068.

Gagnon-Bartsch, Tristan, Matthew Rabin, and Joshua Schwartzstein, "Channeled Attention and Stable Errors," *Working Paper*, 2021.

Gollwitzer, Peter M. and Paschal Sheeran, "Implementation Intentions and Goal Achievement: A Meta-analysis of Effects and Processes," *Advances in Experimental Social Psychology*, 2006, *38*, 69–119.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein, "Learning Through Noticing: Theory and Evidence from a Field Experiment," *Quarterly Journal of Economics*, 2014, *129* (3), 1311–1353.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack, "Unrealistic Expectations and Misguided Learning," *Econometrica*, 2018, *86* (4), 1159–1214.

Hoffman, Mitchell and Stephen V. Burks, "Worker overconfidence: Field evidence and implications for employee turnover and firm profits," *Quantitative Economics*, 2020, *11* (1), 315–348.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman, "Getting to the Top of Mind: How Reminders Increase Saving," *Management Science*, 2016, *62* (12), 3393–3411.

Kaur, Supreet, Sendhil Mullainathan, Suanna Oh, and Frank Schilbach, "Do Financial Concerns Make Workers Less Productive?," *Working Paper no. 28338, National Bureau of Economic Research*, 2021.

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt, "Rational Inattention: A Review," *Journal of Economic Literature*, forthcoming.

Martin, Daniel, "Rational Inattention in Games: Experimental Evidence," *Working Paper*, 2016.

Marx, Benjamin M. and Lesley J. Turner, "Student Loan Nudges: Experimental Evidence on Borrowing and Educational Attainment," *American Economic Journal: Economic Policy*, 2019, *11* (2), 108–141.

Matějka, Filip and Alisdair McKay, "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, January 2015, *105* (1), 272–298.

Milkman, Katherine L., John Beshears, James J. Choi, David Laibson, and Brigitte C. Madrian, "Using implementation intentions prompts to enhance influenza vaccination rates," *Proceedings of the National Academy of Sciences*, 2011, *108* (26), 10415–10420.

Mullainathan, Sendhil, "A Memory-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2002, *117* (3), 735–774.

Nafziger, Julia, "Spillover Effects of Nudges," *Economics Letters*, 2020, *190.*

Nickerson, David W. and Todd Rogers, "Do You Have a Voting Plan? Implementation Intentions, Voter Turnout, and Organic Plan Making," *Psychological Science*, 2010, *21* (2), 194–199.

Oreopoulos, Philip, Richard W. Patterson, Uros Petronijevic, and Nolan G. Pope, "Low-Touch Attempts to Improve Time Management among Traditional and Online College Students," *Journal of Human Resources*, forthcoming.

Shenhav, Amitai, Sebastian Musslick, Falk Lieder, Wouter Kool, Thomas L. Griffiths, Jonathan D. Cohen, and Matthew M. Botvinick, "Toward a Rational and Mechanistic Account of Mental Effort," *Annual Review of Neuroscience*, 2017, *40*, 99–124. PMID: 28375769.

Sims, Christopher A., "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665 – 690.

Tasoff, Joshua and Robert Letzler, "Everyone Believes in Redemption: Nudges and Overoptimism in Costly Task Completion," *Journal of Economic Behavior and Organization*, 2014, *107*, 107–122.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *The Review of Economic Studies*, 2018, *85* (4), 2462–2496.

Wiswall, Matthew and Basit Zafar, "Preference for the Workplace, Investment in Human Capital, and Gender," *The Quarterly Journal of Economics*, 2018, *133* (1), 457–507.

__ and __, "Human Capital Investments and Expectations about Career and Family," *Journal of Political Economy*, 2021, *129* (5), 1361–1424.

Zhang, C. Yiwei, Jeffrey Hemmeter, Judd B. Kessler, Robert D. Metcalfe, and Robert Weathers, "Nudging Timely Wage Reporting: Field Experimental Evidence from the United States Social Supplementary Income Program," *Working Paper no. 27875, National Bureau of Economic Research*, 2021.

Online Appendix

A Mathematical appendix

A.1 Preliminaries

Define the indirect utility functions

$$M_i^1(r) := \max_{\mu} \left\{ \mu B_i(r) - K_i^1(\mu) \right\} \tag{12}$$

$$M_i^0(r) := \max_{\mu} \left\{ \mu B_i(r) - K_i^0(\mu)) \right\} \tag{13}$$

which correspond to the utilities obtained by i if exogenously assigned $j = 1$ or $j = 0$, respectively. Our main assumptions are that

Assumption A1 $B_i(r)$ is smoothly distributed in i for each r.

Assumption A2 $M_i^1(r)$ and $M_i^1(r) - M_i^0(r)$ are smoothly distributed in i for each r.

As one example under which Assumption A2 holds, suppose that types can be partitioned into three-dimensional types $i = (\omega, \eta^1, \eta^2)$, and that $K_i^j = K_\omega^j + \eta^j$ where η^1 and η^2 are random variables on $\mathbb{R}$ that possess differentiable density functions, and that are independent of $G_i(b)$ for each b. In this case, η^j corresponds to the nuisance cost associated with choosing technology j. As another example, suppose that $K_i^0 = K_\omega^0$ and $K_i^1 = \eta K_\omega$, where η is random variable on $\mathbb{R}^+$ with a differentiable density function that is independent of $G_i(b)$. Here, the interpretation is that η captures individual differences in the extent to which $j = 1$ reduces the marginal costs of attention.

Finally, note that the case where $B_i(r) = r$ for all i is a special case of $B_i(r) = \int_{b+r>0}(x+r)g_i(b)db$ where $Pr(b_i = 0) = 1$ for all i.

A.2 Proof of Proposition 1

Proof of part 1

Proof. Individuals choose technology $j = 1$ iff $p \leq M_i^1(r) - M_i^0(r)$. Individual i's WTP for technology $j = 1$ is thus $W_i(r) := M_i^1(r) - M_i^0(r)$. Assumption A1 implies that W_i is

differentiable for each i. Let μ_i^j denote individual i's optimal choice of μ given technology j. Applying the Envelope Theorem to M_i^1 and M_i^0 implies that

$$\begin{aligned}
\frac{d}{dr}\mathbb{E}\left[W_i(r)\right] &= \mathbb{E}\left[\frac{d}{dr}M_i^1(r)\right] - \mathbb{E}\left[\frac{d}{dr}M_i^0(r)\right] \\
&= \mathbb{E}\left[\mu_i^1 \cdot (1-G_i(-r))\right] - \mathbb{E}\left[\mu_i^0 \cdot (1-G_i(-r))\right] && (14) \\
&= Pr(a=1|j=1,r) - Pr(a=1|j=0,r). && (15)
\end{aligned}$$

□

Proof of part 2

Proof. Define

$$V_i(p,r) = \max_{j,\mu} U_i(j,\mu|p,r).$$

Assumption A2 implies that $Pr(j=1|p,r)$ is differentiable in p and r, and that V_i is differentiable in p and r except for a measure zero of i. Thus repeated application of the Envelope Theorem implies that

$$\begin{aligned}
\frac{d}{dp}\frac{d}{dr}\mathbb{E}[V_i(p,r)] &= \frac{d}{dp}\mathbb{E}\left[\frac{d}{dr}V_i(p,r)\right] \\
&= \frac{d}{dp}\left[Pr(j=1|p,r)\mathbb{E}\left[\frac{d}{dr}M_i^1(r)\right] + Pr(j=0|p,r)\mathbb{E}\left[\frac{d}{dr}M_i^0(r)\right]\right] \\
&= \frac{d}{dp}\left[Pr(j=1|p,r)\mathbb{E}\left[\mu_i^1 \cdot (1-G_i(-r))\right] + Pr(j=0|p,r)\mathbb{E}\left[\mu_i^0 \cdot (1-G_i(-r))\right]\right] \\
&= \frac{d}{dp}\left[Pr(j=1|p,r)Pr(a=1|j=1,r) + Pr(j=0|p,r)Pr(a=1|j=0,r)\right] \\
&= \frac{d}{dp}Pr(a=1|p,r)
\end{aligned}$$

where the expectation is taken over all i for which $\frac{d}{dr}V_i(p,r)$ exists (i.e., for all i except possibly those where $M_i^1(r) = M_i^0(r)$).

Similarly applying the Envelope Theorem,

$$\frac{d}{dr}\frac{d}{dp}\mathbb{E}[V_i(p,r)] = \frac{d}{dr}\mathbb{E}\left[\frac{d}{dp}V_i(p,r)\right]$$
$$= -\frac{d}{dr}Pr(j=1|p,r)$$

where the expectation is taken over all i for which $\frac{d}{dp}V_i(p,r)$ exists.

Additionally, note that

$$\begin{aligned}
Pr(a=1|p,r) &= Pr(j=1|p,r)\left(D(a=1|r)+Pr(a=1|j=0,r)\right) \\
&+ Pr(j=0|p,r)Pr(a=1|j=0,r) \\
&= \left(Pr(j=1|p,r)+Pr(j=0|p,r)\right)Pr(a=1|j=0,r) \\
&+ Pr(j=1|p,r)D(a=1|r) \\
&= Pr(a=1|j=0,r)+Pr(j=1|p,r)D(a=1|r).
\end{aligned}$$

Since $Pr(a=1|j=0,r)$ and $D(a=1|r)$ are not functions of p, we thus have that

$$\frac{d}{dp}Pr(a=1|p,r) = \frac{d}{dp}Pr(j=1|p,r)D(a=1|r).$$

Finally, because $\frac{d}{dp}\frac{d}{dr}\mathbb{E}[V_i(p,r)] = \frac{d}{dr}\frac{d}{dp}\mathbb{E}[V_i(p,r)]$, the result follows. □

A.3 Proof of Proposition 2

Proof. Let $\tilde{K}_i^j$ denote the perceived attention production functions. Define the perceived indirect utility functions

$$\tilde{M}_i^1(r) := \max_{\mu}\left\{\mu B_i(r) - \tilde{K}_i^1(\mu)\right\} \tag{16}$$

$$\tilde{M}_i^0(r) := \max_{\mu}\left\{\mu B_i(r) - \tilde{K}_i^0(\mu))\right\} \tag{17}$$

Individuals choose technology $j=1$ iff $p \leq \tilde{M}_i^1(r) - \tilde{M}_i^0(r)$. Individual i's WTP for technology $j=1$ is thus $W_i(r) := \tilde{M}_i^1(r) - \tilde{M}_i^0(r)$. Analogous to the proof of Proposition 1, we have that

$$\frac{d}{dr}W_i(r) = \tilde{Pr}_i(a=1|j=1,r) - \tilde{Pr}_i(a=1|j=0,r)$$

where we continue to use the "tilde" notation to denote subjective beliefs. Taking expectations over i yields the first result.

Similarly, following the proof of the second part of Proposition 1, we have

$$\frac{d}{dp}\frac{d}{dr}\tilde{V}_i(p,r) = \frac{d}{dp}\tilde{Pr}_i(j=1|p,r)\tilde{D}_i(a=1|r) \tag{18}$$

and

$$\frac{d}{dr}\frac{d}{dp}\tilde{V}_i(p,r) = -\frac{d}{dr}\tilde{Pr}_i(j=1|p,r). \tag{19}$$

Taking expectations of (18) and (19) and setting them equal to each other yields the second result. □

A.4 Graphical Illustration

Figure A.1 illustrates the intuition graphically for a representative individual, for the case in which the marginal costs are linear. In this case, the likelihood of executing the task equals the chosen level of attention μ. In analogy to standard theories of competitive supply, individuals' choice of μ with attention technology j is determined by the intersection of the marginal benefit curve $r+b$ and the marginal cost curve $\frac{\partial}{\partial\mu}K^j$. As in theories of competitive supply, the total surplus of an individual with technology $j=0$ at incentive r is equal to the area of triangle OAD less the "fixed cost" $K^0(0)$, which is $(r+b)Pr(a=1|j=0)/2-K^0(0)$. Similarly, the total surplus of an individual with technology $j=1$ is equal to the area of triangle OAF less $K^1(0)$, which is $(r+b)Pr(a=1|j=1)/2-K^1(0)$. Increasing the incentives r by an amount Δ increases surplus by an amount ABCD under technology $j=0$, and by an amount ABEF under technology $j=1$. The change in WTP for technology $j=1$ is thus given by the area DCEF. The area of DCEF is equal to the height, Δ, multiplied by the average of the lengths of DF and CE, which is

$$\left(D(a=1|r)+D(a=1|r+\Delta)\right)/2.$$

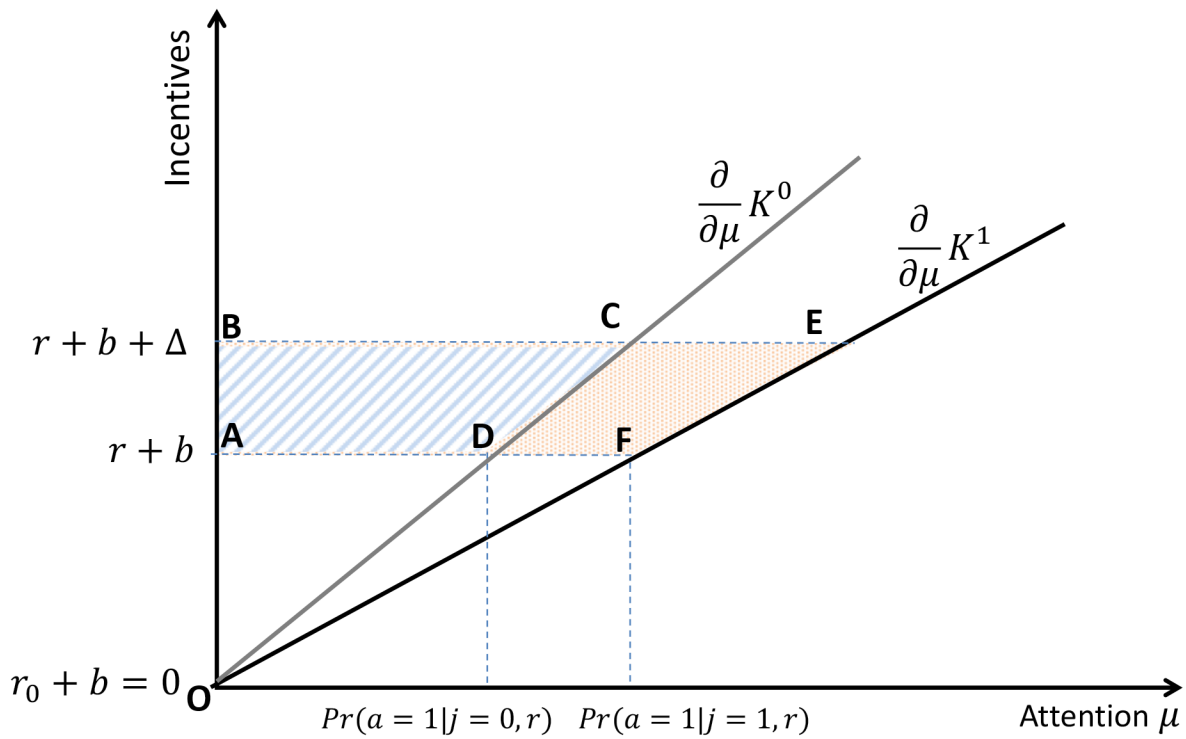
This gives the expression in Corollary 1.

In the limit of very small Δ, the difference between $Pr(a=1|j,r)$ and $Pr(a=1|j,r+\Delta)$ becomes negligible, and thus the area of ABCD can be expressed as

$$\Delta\cdot D(a=1|r)$$

which leads to the first expression in Proposition 1 after dividing by Δ.

Figure A.1: Illustration of Proposition 1



This figure illustrates equation (2) of Proposition 1. The top line (in gray) plots the marginal costs of attention under technology $j=0$, while the bottom line (in black) plots marginal costs under technology $j=1$. The area DCEF corresponds to the change in WTP for technology $j=1$ over $j=0$ when the financial incentive is increased from r to $r+\Delta$.

A.5 Interaction Between Incentives and Reminders

Let $\mu_0^*(r)$ and $\mu_1^*(r)$ be the chosen levels of attention given cost functions K^0 and K^1, respectively, and incentive level r. Let $\Delta\mu^*(r) := \mu_1^*(r) - \mu_0^*(r)$, and suppose that it is non-negative, meaning that the BE increases attentiveness. The impact of the BE on task completion depends on incentives r as follows, where as before G is the CDF of intrinsic benefits b, with PDF g:

$$\begin{aligned}\frac{d}{dr}D(a=1|r) &= \frac{d}{dr}\left[\Delta\mu^*(r)\left(1-G(-r)\right)\right] \\ &= \left(\frac{d}{dr}\Delta\mu^*(r)\right)\left(1-G(-r)\right)+\Delta\mu^*(r)g(-r). \qquad (20)\end{aligned}$$

Now suppose that $\frac{d}{dr}\Delta\mu^*(r) \geq 0$, meaning that either (i) incentives do not affect differences in attentiveness or (ii) incentives actually increase the impact of the BE on attentiveness. In this case, equation (20) is positive, meaning that incentives widen the gap in task completion generated by the BE—that is, higher incentives and the BE are complements. Because $\Delta\mu^*(r)g(-r)$ and $(1-G(-r))$ are positive, for $\frac{d}{dr}D(a=1|r)$ to be negative as in our second experiment, it is necessary that $\frac{d}{dr}\Delta\mu^*(r) < 0$—that is, that the impact of the BE on attentiveness is lower at higher incentives. In particular, since $\mu_j^*(r)$ are non-decreasing in r (under any reasonable model of costly but not necessarily "rational" inattention), this means that the impact of incentives on μ_0^* must be strictly larger than the impact of incentives on μ_1^*, which thus implies that $\frac{d}{dr}\mu_1^*(r) > 0$.

B Additional Results for Experiment 1

Table B.1: Participant Characteristics (Experiment 1)

Students		Alumni	
First-year	0.28	2017	0.22
	(0.45)		(0.41)
Sophomore	0.22	2016	0.18
	(0.41)		(0.39)
Junior	0.23	2015	0.21
	(0.42)		(0.41)
Senior	0.28	2014	0.19
	(0.45)		(0.39)
		2013	0.20
			(0.40)
Female	0.65	Female	0.70
	(0.48)		(0.46)
Male	0.31	Male	0.27
	(0.46)		(0.44)
Non-binary or no answer	0.04	Non-binary or no answer	0.03
	(0.20)		(0.21)
N	686	N	687

This table presents summary statistics for the participants in experiment 1, split between student and alumni groups. These participants were randomized to our various treatments as described in the main text. The Pay-to-Code sample includes 496 participants divided between $2 and $5 incentive arms. The Pay-to-Plan sample includes 487 participants divided between $1 and $2 incentive arms. The remaining participants include 218 control participants and 172 participants assigned to the *Combination* treatment.

Table B.2: The Effect of Coding-Task Incentives on Task Completion

	(1) Week 1	(2) Weeks 1-4	(3) Weeks 1-8
>0	0.036***	0.032***	0.026***
	(0.009)	(0.007)	(0.006)
Obs.	714	714	714
R^2	0.039	0.059	0.064
Control Mean	0.385	0.278	0.210
>10	0.037***	0.034***	0.027***
	(0.009)	(0.007)	(0.006)
Obs.	714	714	714
R^2	0.047	0.067	0.072
Control Mean	0.339	0.243	0.179
>30	0.036***	0.027***	0.023***
	(0.009)	(0.006)	(0.005)
Obs.	714	714	714
R^2	0.043	0.053	0.068
Control Mean	0.239	0.186	0.138
>40	0.038***	0.026***	0.021***
	(0.009)	(0.006)	(0.005)
Obs.	714	714	714
R^2	0.044	0.058	0.074
Control Mean	0.183	0.161	0.119
>50	0.032***	0.022***	0.017***
	(0.008)	(0.005)	(0.004)
Obs.	714	714	714
R^2	0.037	0.058	0.071
Control Mean	0.165	0.142	0.107
>60	0.027***	0.019***	0.013***
	(0.008)	(0.005)	(0.004)
Obs.	714	714	714
R^2	0.044	0.066	0.066
Control Mean	0.138	0.118	0.093
Controls	Yes	Yes	Yes
Campus × Student FE	Yes	Yes	Yes

This table presents estimates for the effect of coding-task incentives (in dollars) on task completion. Each panel of the table corresponds to an analysis of whether participants completed at least that number of minutes of the coding task in a given week. The columns correspond to different periods during the experiment over which the effect of the incentives is tested: Column (1) shows the effect in week 1, Column (2) shows the effect for weeks 1-4, and Column (3) shows the effect over all weeks. In Column (1), the dependent variable is an indicator for whether a participant completed at least that many minutes of the coding task in the first week. In Columns (2) and (3), the dependent variable is the mean of the indicators, constructed as in Column (1), for each of the weeks being considered. Each panel-by-column corresponds to a separate specification, and thus 18 distinct specifications are shown in the table. Standard errors are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table B.3: The Effect of Plan-Making Incentives on Task Completion

	(1) Week 1	(2) Weeks 1-4	(3) Weeks 1-8
>0	0.037*	0.029*	0.014
	(0.022)	(0.015)	(0.012)
Obs.	705	705	705
R^2	0.041	0.040	0.046
Control Mean	0.385	0.278	0.210
>10	0.037*	0.027*	0.014
	(0.021)	(0.014)	(0.011)
Obs.	705	705	705
R^2	0.045	0.045	0.046
Control Mean	0.339	0.243	0.179
>30	0.045**	0.023*	0.010
	(0.020)	(0.013)	(0.010)
Obs.	705	705	705
R^2	0.054	0.042	0.045
Control Mean	0.239	0.186	0.138
>40	0.036**	0.019	0.008
	(0.018)	(0.012)	(0.009)
Obs.	705	705	705
R^2	0.034	0.036	0.041
Control Mean	0.183	0.161	0.119
>50	0.034*	0.015	0.005
	(0.018)	(0.011)	(0.008)
Obs.	705	705	705
R^2	0.035	0.039	0.042
Control Mean	0.165	0.142	0.107
>60	0.027*	0.013	0.002
	(0.016)	(0.010)	(0.008)
Obs.	705	705	705
R^2	0.044	0.038	0.042
Control Mean	0.138	0.118	0.093
Controls	Yes	Yes	Yes
Campus × Student FE	Yes	Yes	Yes

This table presents estimates for the effect of plan-making incentives (in dollars) on task completion. Each panel of the table corresponds to an analysis of whether participants completed at least that number of minutes of the task in a given week. The columns correspond to different periods during the experiment over which the effect of the incentives is tested: Column (1) shows the effect in week 1, Column (2) shows the effect for weeks 1-4, and Column (3) shows the effect for all weeks. In Column (1), the dependent variable is an indicator for whether a participant completed at least that many minutes of the coding task in the first week. In Columns (2) and (3), the dependent variable is the mean of the indicators, constructed as in Column (1), for each of the weeks being considered. Each panel-by-column corresponds to a separate specification, and thus 18 distinct specifications are shown in the table. Standard errors are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table B.4: The Effect of Plan-Making Incentives on Task Completion (2SLS)

A. The Effect on Plan Making (First Stage)

	(1) Week 1	(2) Weeks 1-4	(3) Weeks 1-8
$1 Plan	0.282***	0.285***	0.240***
	(0.048)	(0.033)	(0.030)
$2 Plan	0.368***	0.297***	0.242***
	(0.033)	(0.028)	(0.025)
Obs.	705	705	705
R^2	0.144	0.189	0.157
Control Mean	0.381	0.150	0.082
Controls	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes

B. The Effect on Coding Task Completion (Reduced Form)

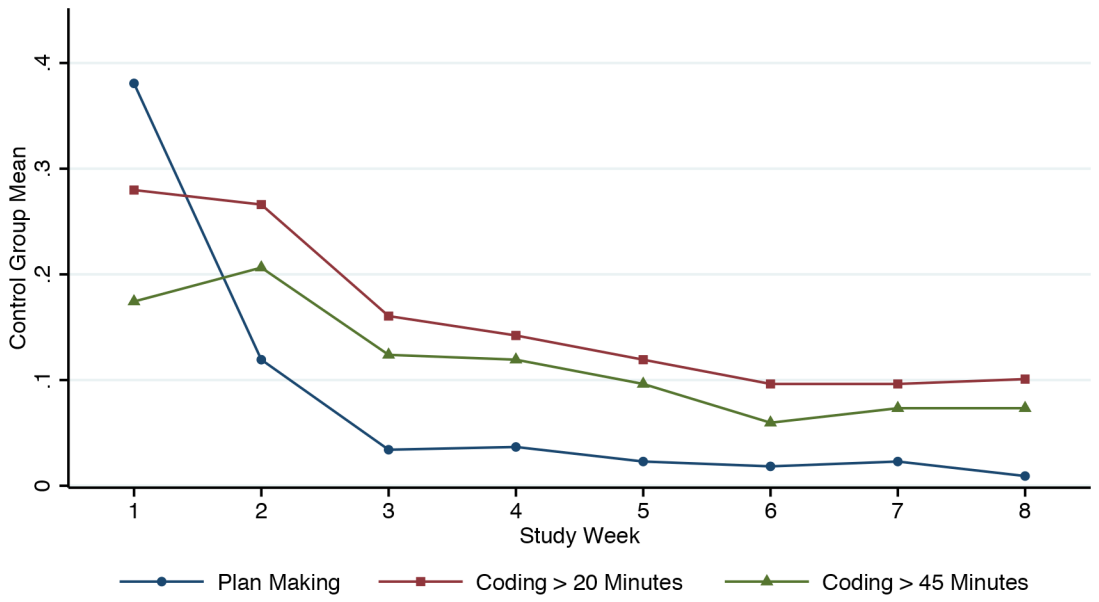
	(1) >20 (1)	(2) >20 (1-4)	(3) >20 (1-8)	(4) >45 (1)	(5) >45 (1-4)	(6) >45 (1-8)
$1 Plan	0.034	0.017	0.013	0.034	-0.000	0.005
	(0.048)	(0.032)	(0.025)	(0.043)	(0.027)	(0.021)
$2 Plan	0.079*	0.054**	0.026	0.076**	0.032	0.012
	(0.040)	(0.027)	(0.021)	(0.036)	(0.023)	(0.018)
Obs.	705	705	705	705	705	705
R^2	0.057	0.049	0.051	0.036	0.035	0.041
Control Mean	0.280	0.212	0.158	0.174	0.156	0.116
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes	Yes	Yes	Yes

C. The Effect of Plan Making on Coding Task Completion (IV)

	(1) >20 (1)	(2) >20 (1-4)	(3) >20 (1-8)	(4) >45 (1)	(5) >45 (1-4)	(6) >45 (1-8)
Plan Making	0.203**	0.146*	0.092	0.197**	0.076	0.041
	(0.102)	(0.080)	(0.078)	(0.093)	(0.070)	(0.066)
Obs.	705	705	705	705	705	705
R^2	0.143	0.151	0.120	0.091	0.094	0.076
Control Mean	0.280	0.212	0.158	0.174	0.156	0.116
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Campus FE	Yes	Yes	Yes	Yes	Yes	Yes

This table shows estimates for the effect of plan-making incentives on plan making and task completion using treatment dummies rather than a linear plan-making incentive variable. Panel A shows estimates of the effect of plan-making incentives on whether or not participants made a plan. Column (1) shows the effect of plan-making incentives in week 1. Column (2) shows the average effect over weeks 1-4. Column (3) shows the average effect over all weeks. Panel B shows the effect of plan-making incentives on task completion. Columns (1)–(3) show the effect on an indicator variable for whether or not the participant worked on the coding task for more than 20 minutes: Column (1) estimates the effect over week 1, Column (2) over weeks 1-4, and Column (3) over all weeks. Columns (4)–(6) show analogous estimates, but for an indicator variable for whether or not the participant worked on the task for more than 45 minutes each week. Panel C shows the 2SLS estimates instrumenting for whether or not participants made a plan using the plan-making treatment dummies as instruments. The dependent variables are the same as those in Panel B. Standard errors are shown in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Figure B.1: Experiment 1 Control Group Means (Week-by-Week)



This figure shows control group means for plan making and completing at least 20 minutes or at least 45 minutes of the coding course for each week of the study.

C Additional Results for Experiment 2

Table C.1: The Effect of Incentive, Delay, and Reminders on Part-2 Survey Completion (Categorical Delay)

	Completed Part-2 Survey		
	(1)	(2)	(3)
Received Reminder	0.23***	0.15***	0.16***
	(0.021)	(0.045)	(0.044)
High Incentive	0.08***	0.14***	0.13***
	(0.021)	(0.029)	(0.031)
1-Week Delay	-0.16***	-0.22***	-0.22***
	(0.029)	(0.041)	(0.072)
3-Week Delay	-0.17***	-0.27***	-0.27***
	(0.030)	(0.042)	(0.067)
6-Week Delay	-0.21***	-0.32***	-0.30***
	(0.030)	(0.041)	(0.063)
Received Reminder × High Incentive		-0.13***	-0.11***
		(0.042)	(0.040)
1-Week Delay × Received Reminder		0.13**	0.10
		(0.058)	(0.075)
3-Week Delay × Received Reminder		0.20***	0.20***
		(0.060)	(0.057)
6-Week Delay × Received Reminder		0.22***	0.18***
		(0.059)	(0.063)
Constant	0.55***	0.59***	0.59***
	(0.025)	(0.032)	(0.043)
Observations	2,076	2,076	2,076
Number of Participants	2,076	2,076	2,076
S.E. Clustered by P1 & P2 Date			X
P1 Date FE			X

This table estimates how survey completion varies with reminders, delay, and whether participants are offered high incentives (i.e., \$11 or \$12) or low incentives (i.e., \$3 or \$4) to complete the survey. The 2-day delay variable is omitted so the 2-day delay is the excluded group. This table only includes participants who were randomly assigned to receive or not receive reminders. Column (3) reproduces Column (2) with fixed effects for the date that part 1 of the study was taken and with standard errors clustered for the date the participant completed part 1 and the date part 2 was made available to them. Standard errors are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.2: The Effect of Incentive and Delay on Willingness to Pay for Reminders (Only Participants Randomized for Reminders)

	WTP for Reminders ($)					
	(1)	(2)	(3)	(4)	(5)	(6)
Extra $1	0.07***	0.08	0.06	0.07***	0.10***	0.11*
	(0.018)	(0.055)	(0.038)	(0.018)	(0.019)	(0.055)
High Incentive	0.96***	0.96***	1.14***	0.96***	0.96***	0.96***
	(0.082)	(0.082)	(0.118)	(0.082)	(0.082)	(0.082)
Extra $1 × High Incentive	-0.05	-0.05	-0.01	-0.05	-0.06	-0.06
	(0.050)	(0.050)	(0.116)	(0.050)	(0.051)	(0.051)
Ln(P2 Delay)		-0.07***	-0.04***			-0.07***
		(0.022)	(0.013)			(0.023)
Extra $1 × Ln(P2 Delay)		-0.00	0.01			-0.00
		(0.022)	(0.014)			(0.022)
High Incentive × Ln(P2 Delay)			-0.08**			
			(0.036)			
Extra $1 × Ln(P2 Delay)			-0.02			
× High Incentive			(0.044)			
Constant	0.51***	0.68***	0.59***	0.51***	0.51***	0.68***
	(0.034)	(0.060)	(0.044)	(0.034)	(0.034)	(0.060)
Observations	33,216	33,216	33,216	33,216	33,216	33,216
Number of Participants	2,076	2,076	2,076	2,076	2,076	2,076
Specification	OLS	OLS	OLS	OLS	Tobit	Tobit
P1 Date FE				X		

This table estimates how willingness to pay for reminders varies with the natural log of delay (in days) and incentives to complete the survey. This table only includes participants who were randomly assigned to receive or not receive reminders. The High Incentive variable is an indicator for being asked about an incentive of $11 or $12. The Extra $1 variable is an indicator for being asked about an incentive of $4 or $12. Column (4) reproduces Column (1) with fixed effects for the date that part 1 of the survey was taken; Columns (5) and (6) reproduce Columns (1) and (2) using Tobit estimates with censors at -$4 and $4 for the low-incentive group and censors at -$12 and $12 for the high-incentive group. Standard errors, clustered at the participant level, are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.3: The Effect of Incentive and Delay on Willingness to Pay for Reminders by MPL

	WTP for Reminders ($)			
	(1)	(2)	(3)	(4)
Extra $1	0.15*	0.08	0.11***	0.09***
	(0.078)	(0.050)	(0.034)	(0.026)
High Incentive	0.58***	0.68***	0.79***	0.77***
	(0.193)	(0.154)	(0.143)	(0.139)
Extra $1 × High Incentive	0.18	0.02	-0.10	-0.02
	(0.244)	(0.150)	(0.106)	(0.077)
Constant	0.45***	0.49***	0.50***	0.51***
	(0.063)	(0.053)	(0.048)	(0.046)
Observations	4,612	9,224	13,836	18,448
Number of Participants	2,306	2,306	2,306	2,306
First T MPLs	$T = 2$	$T = 4$	$T = 6$	$T = 8$

This table estimates how willingness to pay for reminders varies with incentives to complete the survey. The High Incentive variable is an indicator for being asked about an incentive of $11 or $12. The Extra $1 variable is an indicator for being asked about an incentive of $4 or $12. Column (1) shows these estimates when limited to the first 2 MPLs participants are asked about, Column (2) shows these estimates when limited to the first 4 MPLs, Column (3) shows these estimates when limited to the first 6 MPLs, and Column (4) shows these estimates when limited to the first 8 MPLs. Standard errors, clustered at the participant level, are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.4: The Effect of Incentive and Delay on WTP for Reminders (Categorical Delay)

	WTP for Reminders ($)				
	(1)	(2)	(3)	(4)	(5)
Extra $1	0.07***	0.05	0.07***	0.10***	0.08
	(0.017)	(0.047)	(0.017)	(0.018)	(0.047)
High Incentive	0.93***	0.93***	0.93***	0.93***	0.93***
	(0.077)	(0.077)	(0.077)	(0.077)	(0.077)
Extra $1 × High Incentive	-0.06	-0.06	-0.06	-0.07	-0.07
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
1-Week Delay		-0.16***			-0.16***
		(0.054)			(0.055)
3-Week Delay		-0.19***			-0.19***
		(0.062)			(0.062)
6-Week Delay		-0.25***			-0.25***
		(0.067)			(0.067)
1-Week Delay × Extra $1		0.03			0.03
		(0.073)			(0.074)
3-Week Delay × Extra $1		0.02			0.02
		(0.068)			(0.068)
6-Week Delay × Extra $1		0.03			0.03
		(0.068)			(0.068)
Constant	0.51***	0.66***	0.51***	0.50***	0.65***
	(0.032)	(0.047)	(0.032)	(0.032)	(0.047)
Observations	36,896	36,896	36,896	36,896	36,896
Number of Participants	2,306	2,306	2,306	2,306	2,306
Specification	OLS	OLS	OLS	Tobit	Tobit
P1 Date FE			X		

This table estimates how willingness to pay for reminders varies with incentives to complete the survey. The High Incentive variable is an indicator for being asked about an incentive of $11 or $12. The Extra $1 variable is an indicator for being asked about an incentive of $4 or $12. Column (2) maintains the specification in Column (1) and adds controls for delay; Column (3) shows Column (1) with fixed effects for the date that part 1 of the survey was taken; Columns (4) and (5) reproduce Columns (1) and (2) using Tobit estimates with censors at -$4 and $4 for the low-incentive group and censors at -$12 and $12 for the high-incentive group. Standard errors, clustered at the participant level, are shown in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

D Additional Results for Experiment 3

Table D.1: Replication of Table 6 with Tobit Models

	Willingness to Pay ($)					
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive ($)	0.11***	-0.00	0.05	0.03	0.08*	0.03
	(0.035)	(0.031)	(0.046)	(0.042)	(0.052)	(0.073)
Incentive ($) × Block 2	-0.07*	0.04	0.01	0.04	0.00	0.02
	(0.038)	(0.037)	(0.051)	(0.052)	(0.065)	(0.076)
Incentive ($) × Feedback			0.11	-0.07	0.12	0.11
			(0.069)	(0.061)	(0.093)	(0.100)
Incentive ($) × Block 2			-0.16**	0.01	-0.22**	-0.11
× Feedback			(0.076)	(0.075)	(0.103)	(0.110)
Block 2	0.14	-0.18	-0.10	-0.21	-0.17	-0.06
	(0.126)	(0.124)	(0.178)	(0.177)	(0.236)	(0.260)
Block 2 × Feedback			0.49*	0.05	0.89***	0.17
			(0.251)	(0.248)	(0.343)	(0.359)
Feedback			-0.23	0.22	-0.37	-0.12
			(0.250)	(0.214)	(0.347)	(0.354)
Constant	0.17	0.51***	0.29*	0.40***	0.28	0.30
	(0.125)	(0.107)	(0.172)	(0.148)	(0.212)	(0.260)
Observations	3,996	4,794	3,996	4,794	1,788	2,208
Number of Participants	666	799	666	799	298	368
Participant B1 Acc. Diff.	All	All	All	All	≤ 0	> 0
Arm	Length	Discernibility	Length	Discernibility	Length	Length

This table replicates Table 6 but presents Tobit estimates with censors at 4 and -4 for the effect of accuracy incentive, block order, and whether the participant received feedback on their block 1-performance on willingness to pay for an easy task (i.e., a task with shorter length in the length arm, or a task with increased discernibility in the discernibility arm). Standard errors, clustered at the participant level, are shown in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table D.2: Replication of Table 6, Dropping the 10% Fastest Participants

	Willingness to Pay ($)					
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive ($)	0.10***	-0.01	0.04	0.02	0.08	0.01
	(0.037)	(0.033)	(0.051)	(0.046)	(0.061)	(0.076)
Incentive ($) × Block 2	-0.07*	0.03	0.02	0.03	0.03	0.02
	(0.041)	(0.040)	(0.055)	(0.056)	(0.075)	(0.078)
Incentive ($) × Feedback			0.12*	-0.07	0.14	0.12
			(0.073)	(0.066)	(0.096)	(0.103)
Incentive ($) × Block 2			-0.19**	0.00	-0.31***	-0.12
× Feedback			(0.082)	(0.081)	(0.117)	(0.112)
Block 2	0.16	-0.12	-0.13	-0.15	-0.24	-0.06
	(0.138)	(0.133)	(0.198)	(0.189)	(0.277)	(0.275)
Block 2 × Feedback			0.59**	0.08	1.24***	0.20
			(0.274)	(0.266)	(0.397)	(0.371)
Feedback			-0.15	0.17	-0.30	-0.06
			(0.270)	(0.231)	(0.389)	(0.369)
Constant	0.17	0.50***	0.24	0.42***	0.18	0.28
	(0.135)	(0.115)	(0.192)	(0.160)	(0.252)	(0.275)
Observations	3,438	4,314	3,438	4,314	1,344	2,094
Number of Participants	573	719	573	719	224	349
Participant B1 Acc. Diff.	All	All	All	All	≤ 0	> 0
Arm	Length	Discernibility	Length	Discernibility	Length	Length

This table replicates Table 6 on the effect of accuracy incentive, block order, and whether the participant received feedback on their block 1-performance on willingness to pay for an easy task (i.e., a task with shorter length in the length arm, or a task with increased discernibility in the discernibility arm) after dropping participants in the top 10% of fastest task times in the length arm and the top 10% of fastest task times in the discernibility arm. Standard errors, clustered at the participant level, are shown in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table D.3: The Effect of Block and Feedback on Block Accuracy Difference

	Accuracy Difference Between Easy and Baseline Tasks					
	(1)	(2)	(3)	(4)	(5)	(6)
Block 2	0.03	0.01	0.20***	-0.13***	0.19***	-0.10***
	(0.021)	(0.018)	(0.018)	(0.020)	(0.027)	(0.029)
Feedback	0.03	-0.01			0.02	0.03
	(0.023)	(0.020)			(0.016)	(0.023)
Block 2 × Feedback	-0.04	-0.01			0.00	-0.07*
	(0.030)	(0.025)			(0.036)	(0.039)
Constant	0.18***	0.26***	-0.06***	0.40***	-0.07***	0.39***
	(0.016)	(0.014)	(0.008)	(0.012)	(0.013)	(0.016)
Observations	1,332	1,598	596	736	596	736
Number of Participants	666	799	298	368	298	368
Participant B1 Acc. Diff.	All	All	≤ 0	> 0	≤ 0	> 0
Arm	Length	Discernibility	Length	Length	Length	Length

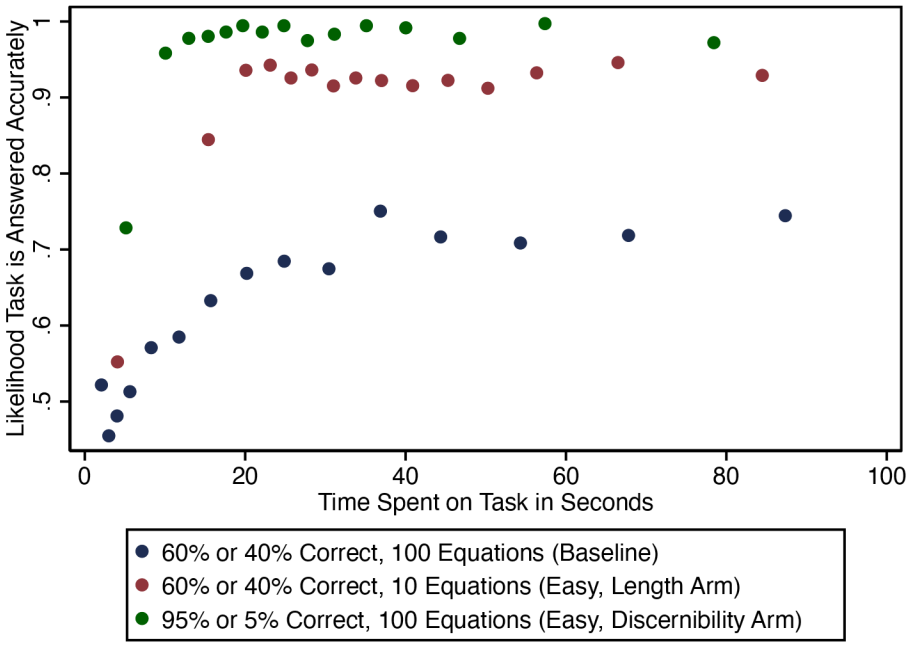
This table estimates the effect of block order and feedback on the accuracy difference between easy and baseline tasks within a block. The accuracy difference is constructed by taking the difference between the percentage of easy tasks answered correctly and the percentage of baseline tasks answered correctly in a block. Column (1) shows OLS estimates for participants in the length arm; Column (2) shows OLS estimates for participants in the discernibility arm; Columns (3) and (5) restrict to participants in the length arm who had a block-1 accuracy difference less than or equal to 0 (i.e., who were at least as accurate in the baseline tasks as in the easy tasks); Columns (4) and (6) restrict to participants in the length arm who had a block-1 accuracy difference greater than 0 (i.e., who were more accurate in the easy tasks than the baseline tasks). Standard errors, clustered at the participant level, are shown in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table D.4: The Effect of Block, Feedback, and Accuracy Difference on Within-Block Time Spent on Baseline vs. Easy Tasks

	Average Difference in Time Spent on Baseline vs. Easy Tasks, By Block					
	(1)	(2)	(3)	(4)	(5)	(6)
Block 2	-21.72***	-13.78**	-19.24***	-11.59**	-16.63***	-9.04*
	(4.125)	(6.078)	(3.562)	(5.444)	(3.144)	(4.740)
B1 Acc. Diff ≤ 0	115.45***	106.94***	103.16***	93.86***	79.81***	74.50***
	(10.006)	(14.582)	(8.194)	(11.926)	(6.141)	(9.036)
Block 2 $\times$ B1 Acc. Diff ≤ 0	-43.62***	-54.99***	-33.48***	-42.74***	-16.53***	-26.72***
	(7.547)	(11.468)	(5.928)	(8.959)	(4.665)	(6.987)
Feedback		3.53		3.40		6.12
		(9.958)		(8.729)		(7.343)
Block 2 $\times$ Feedback		-16.07*		-15.47**		-15.34**
		(8.211)		(7.074)		(6.236)
B1 Acc. Diff $\leq 0 \times$ Feedback		17.26		18.86		10.78
		(19.988)		(16.347)		(12.243)
Block 2 $\times$ B1 Acc. Diff ≤ 0		23.01		18.74		20.61**
$\times$ Feedback		(15.056)		(11.823)		(9.290)
Constant	19.23***	17.48**	16.51***	14.83**	11.79***	8.76
	(4.978)	(7.384)	(4.366)	(6.615)	(3.673)	(5.423)
Observations	1,332	1,332	1,332	1,332	1,332	1,332
Number of Participants	666	666	666	666	666	666
Winsorized at T Seconds	No	No	$T = 300$	$T = 300$	$T = 180$	$T = 180$

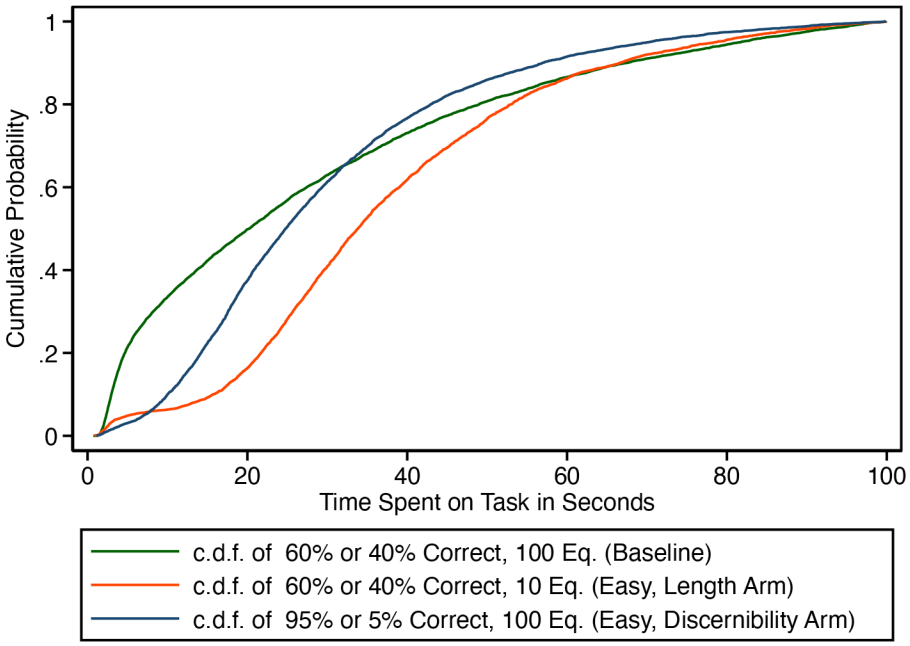
This table estimates the effect of block, feedback and accuracy difference on the average difference in time spent on baseline vs. easy tasks for participants in the length arm. The dependent variable is constructed by taking the difference between average time spent on the baseline tasks in a block and the average time spent on the easy tasks in the same block. By-block accuracy difference is constructed by taking the difference between the percentage of easy tasks answered correctly and the percentage of baseline tasks answered correctly in a block. Column (3) maintains the specification in Column (1) and winsorizes at 300 seconds; Column (4) maintains the specification in Column (2) and winsorizes at 300 seconds; Column (5) maintains the specification in Column (1) and winsorizes at 180 seconds; Column (6) maintains the specification in Column (2) and winsorizes at 180 seconds. Standard errors, clustered at the participant level, are shown in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Figure D.1: The Likelihood of Answering Tasks Accurately by Time Spent, Dropping Observations > 100 Seconds



This figure includes binned scatterplots that displays how accuracy varies with the time spent on the three types of tasks in seconds after dropping responses in which a participant spent more than 100 seconds on a task. Here, observations have been separated into 15 equal-sized bins in each binned scatterplot.

Figure D.2: CDFs of Time Spent on Task by Length and Discernibility (Dropping Observations > 100 Seconds)



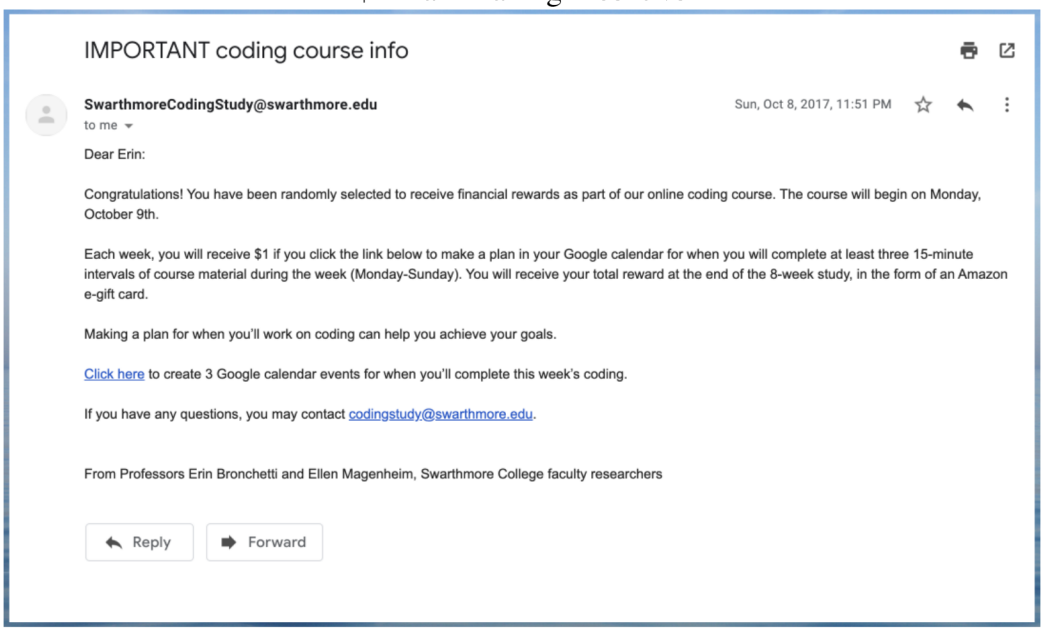
This figure displays the CDFs of time spent on a task in seconds by task type after dropping responses in which a participant spent more than 100 seconds on a task.

E Screenshots Appendix

E.1 Experiment 1

Figure E.1: *Pay-to-Plan* Treatment Emails, Week 1

$1 Plan-Making Incentive



IMPORTANT coding course info

SwarthmoreCodingStudy@swarthmore.edu Sun, Oct 8, 2017, 11:51 PM
to me

Dear Erin:

Congratulations! You have been randomly selected to receive financial rewards as part of our online coding course. The course will begin on Monday, October 9th.

Each week, you will receive $1 if you click the link below to make a plan in your Google calendar for when you will complete at least three 15-minute intervals of course material during the week (Monday-Sunday). You will receive your total reward at the end of the 8-week study, in the form of an Amazon e-gift card.

Making a plan for when you'll work on coding can help you achieve your goals.

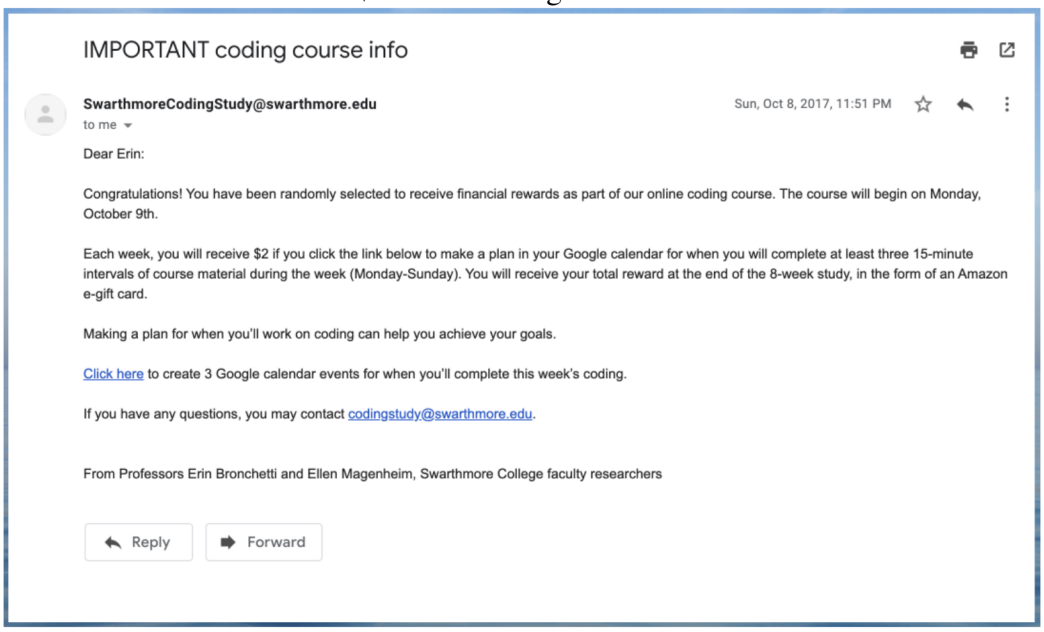
Click here to create 3 Google calendar events for when you'll complete this week's coding.

If you have any questions, you may contact codingstudy@swarthmore.edu.

From Professors Erin Bronchetti and Ellen Magenheim, Swarthmore College faculty researchers

Reply Forward

$2 Plan-Making Incentive



IMPORTANT coding course info

SwarthmoreCodingStudy@swarthmore.edu Sun, Oct 8, 2017, 11:51 PM
to me

Dear Erin:

Congratulations! You have been randomly selected to receive financial rewards as part of our online coding course. The course will begin on Monday, October 9th.

Each week, you will receive $2 if you click the link below to make a plan in your Google calendar for when you will complete at least three 15-minute intervals of course material during the week (Monday-Sunday). You will receive your total reward at the end of the 8-week study, in the form of an Amazon e-gift card.

Making a plan for when you'll work on coding can help you achieve your goals.

Click here to create 3 Google calendar events for when you'll complete this week's coding.

If you have any questions, you may contact codingstudy@swarthmore.edu.

From Professors Erin Bronchetti and Ellen Magenheim, Swarthmore College faculty researchers

Reply Forward

Figure E.2: *Pay-to-Code* Treatment Emails, Week 1

$2 Coding-Task Incentive

IMPORTANT coding course info

SwarthmoreCodingStudy@swarthmore.edu Sun, Oct 8, 2017, 11:50 PM
to me

Dear Erin:

Congratulations! You have been randomly selected to receive financial rewards for completing parts of the online coding course. The course will begin on Monday, October 9th.

Each week, you will receive $2 if you complete at least three 15-minute intervals of course material during that week (Monday-Sunday). You will receive your total reward at the end of the 8-week study, in the form of an Amazon e-gift card.

Making a plan for when you'll work on coding can help you achieve your goals.

Click here to create 3 Google calendar events for when you'll complete this week's coding.

If you have any questions, you may contact codingstudy@swarthmore.edu.

From Professors Erin Bronchetti and Ellen Magenheim, Swarthmore College faculty researchers

Reply Forward

$5 Coding-Task Incentive

IMPORTANT coding course info

SwarthmoreCodingStudy@swarthmore.edu Sun, Oct 8, 2017, 11:50 PM
to me

Dear Erin:

Congratulations! You have been randomly selected to receive financial rewards for completing parts of the online coding course. The course will begin on Monday, October 9th.

Each week, you will receive $5 if you complete at least three 15-minute intervals of course material during that week (Monday-Sunday). You will receive your total reward at the end of the 8-week study, in the form of an Amazon e-gift card.

Making a plan for when you'll work on coding can help you achieve your goals.

Click here to create 3 Google calendar events for when you'll complete this week's coding.

If you have any questions, you may contact codingstudy@swarthmore.edu.

From Professors Erin Bronchetti and Ellen Magenheim, Swarthmore College faculty researchers

Reply Forward

Figure E.3: *Combined* and Control Group Emails, Week 1

Combined Treatment ($1 Plan-Making and $2 Coding-Task Incentive)

IMPORTANT coding course info

SwarthmoreCodingStudy@swarthmore.edu Sun, Oct 8, 2017, 11:51 PM

to me

Dear Erin:

Congratulations! You have been randomly selected to receive financial rewards as part of our online coding course. The course will begin on Monday, October 9th.

Each week, you will receive $1 if you click the link below to make a plan in your Google calendar for when you will complete at least three 15-minute intervals of course material during the week (Monday-Sunday).

Each week, you will also receive $2 if you complete at least three 15-minute intervals of course material during that week (Monday-Sunday).

You will receive your total reward at the end of the 8-week study, in the form of an Amazon e-gift card.

Making a plan for when you'll work on coding can help you achieve your goals.

Click here to create 3 Google calendar events for when you'll complete this week's coding.

If you have any questions, you may contact codingstudy@swarthmore.edu.

From Professors Erin Bronchetti and Ellen Magenheim, Swarthmore College faculty researchers

Reply Forward

Control Group

IMPORTANT coding course info

SwarthmoreCodingStudy@swarthmore.edu Sun, Oct 8, 2017, 11:29 PM

to me

Dear Erin:

Congratulations on enrolling in our free, 8-week coding course in HTML, Java, and WebDev! The course will begin on Monday, October 9th.

Each week, we encourage you to complete at least three 15-minute intervals of course material sometime during the week. Making a plan for when you'll work on coding can help you achieve your goals for the course.

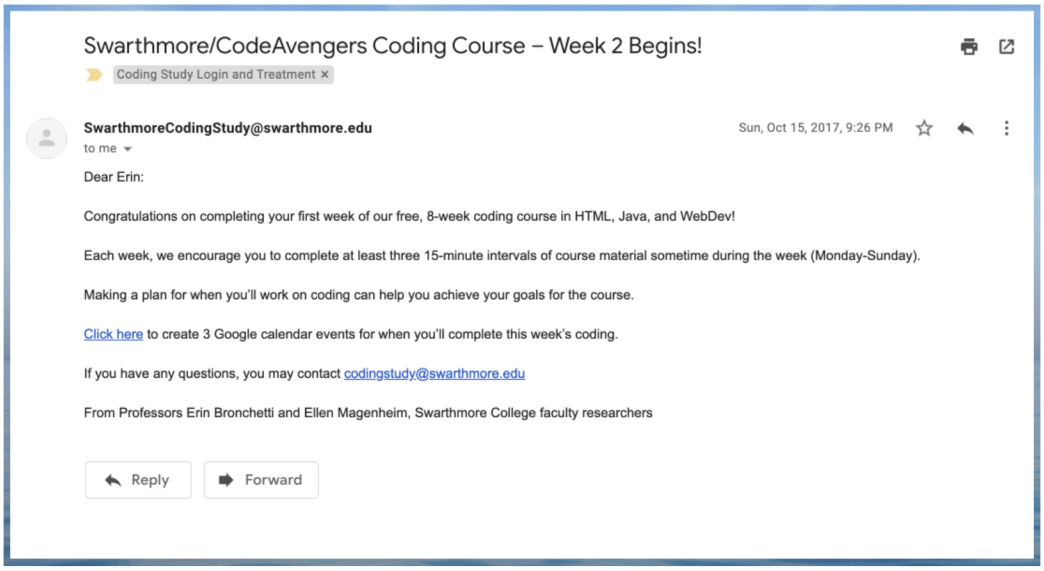
Click here to create 3 Google calendar events for when you'll complete this week's coding.

If you have any questions, you may contact codingstudy@swarthmore.edu.

From Professors Erin Bronchetti and Ellen Magenheim, Swarthmore College faculty researchers

Reply Forward

Figure E.4: Weekly Reminder Email, All Groups



Swarthmore/CodeAvengers Coding Course – Week 2 Begins!

Coding Study Login and Treatment ×

SwarthmoreCodingStudy@swarthmore.edu Sun, Oct 15, 2017, 9:26 PM

to me

Dear Erin:

Congratulations on completing your first week of our free, 8-week coding course in HTML, Java, and WebDev!

Each week, we encourage you to complete at least three 15-minute intervals of course material sometime during the week (Monday-Sunday).

Making a plan for when you'll work on coding can help you achieve your goals for the course.

Click here to create 3 Google calendar events for when you'll complete this week's coding.

If you have any questions, you may contact codingstudy@swarthmore.edu

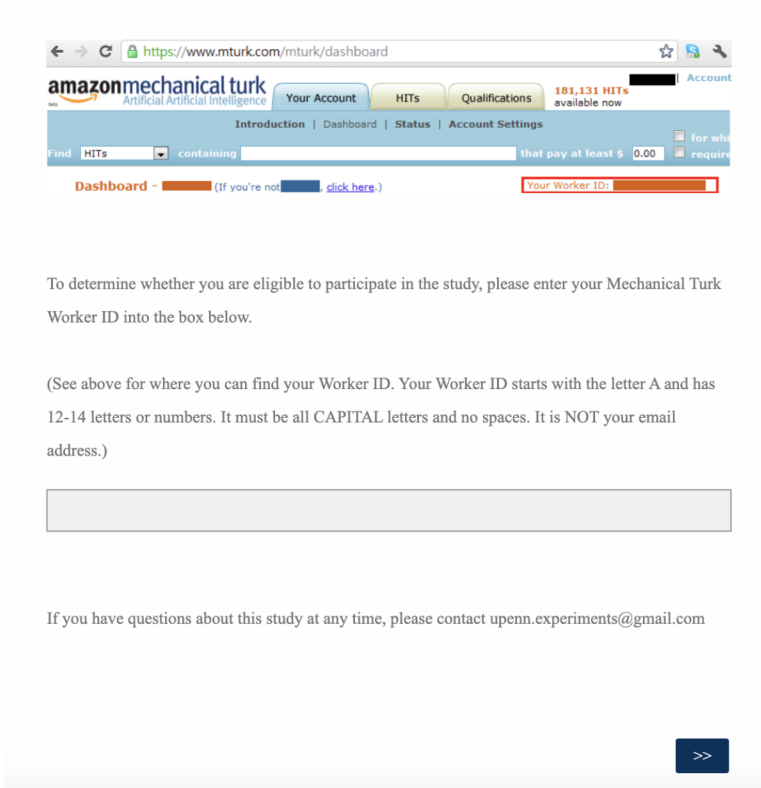
From Professors Erin Bronchetti and Ellen Magenheim, Swarthmore College faculty researchers

Reply Forward

E.2 Experiment 2

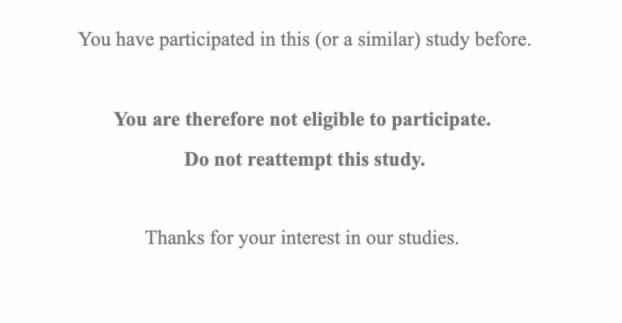
E.2.1 Part 1

Figure E.5: Eligibility Screen



This is the screen in which the participant entered their MTurk ID.

Figure E.6: If Ineligible Screen



If participants had participated in the study at an earlier date, they were shown this screen and excluded from participating.

Figure E.7: Consent Form

This is a consent form. Please read and click below to continue.

Study Background and Purpose: This study examines decision-making. Your participation in this research will take approximately 15 minutes today and you will have the option to complete an additional task for approximately 20 minutes during a one-week window between Sunday, October 3rd and Friday, November 19th.

What Happens in this Research Study: If you decide to participate, you will answer questions.

Payments: There are no known costs to you for participating in this research study except for your time. Upon completion of the survey, you will be given a code that you can submit to MTurk so that you can receive your payment. You will be paid $2.50 for completing the entire survey today, and you will have the possibility of earning an additional bonus payment.

Confidentiality: Your data will be anonymous and will not be linked to your identity.

Voluntary Participation: Participating in this research is voluntary. You can withdraw from the study at any time.

Contact: If you have questions, concerns, or complaints regarding this research, please contact the researcher at upenn.experiments@gmail.com. If a member of the research team cannot be reached or you want to talk to someone other than those working on the study, you may contact the Office of Regulatory Affairs with any question, concerns or complaints at the University of Pennsylvania by calling (215) 898-2614.

Agreement to Participate: By clicking to continue, you are indicating that you have read this consent form and that you voluntarily agree to participate in the study.

>>

Figure E.8: Attention Check (first attempt)

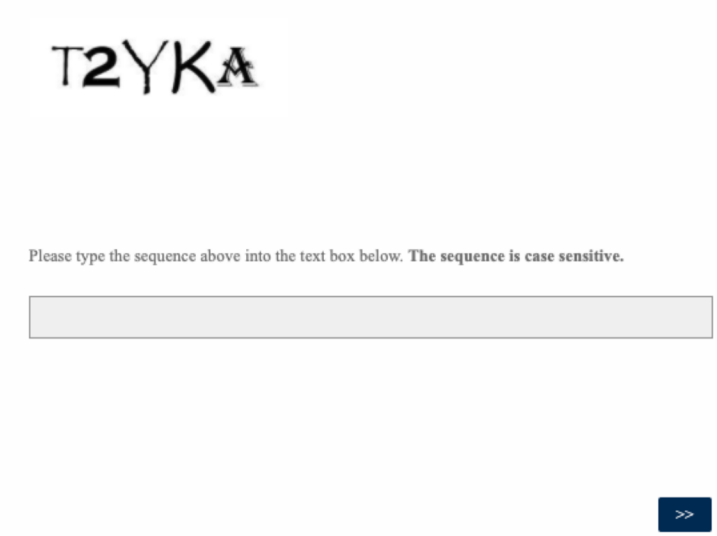
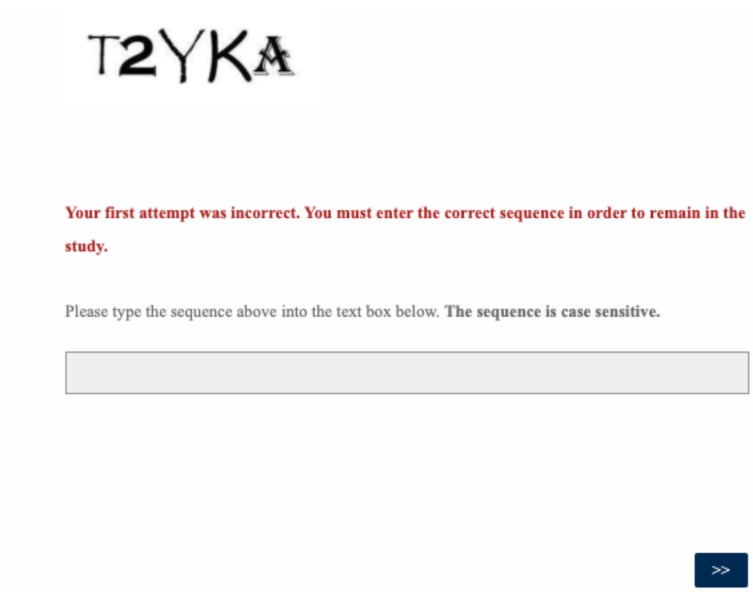


Figure E.9: Attention Check (second attempt)



If participants answered the attention check question incorrectly the first time, they saw this screen which warned them that failure to enter the sequence correctly would remove them from the study.

Figure E.10: Instructions, Screen 1

Thank you for participating in this HIT. This HIT has two parts. Part 1 of the study should take 15 minutes to complete right now. **For completing part 1 of the study, you will earn $2.50. You may also earn a "part 1 bonus" payment.**

In part 2 of the study, you will be asked to complete a survey that will take you about 20 minutes and must be completed in one sitting. **If you complete part 2 of the study, you will earn a "part 2 bonus" payment.** You cannot complete part 2 of the study now. You can only complete part 2 at some point in the future. (We will explain the details of when you can complete part 2 of the study on the next screen.)

In part 2 of the study, we will ask you **40 hypothetical questions** about whether or not you would take a particular gamble. These questions are **hypothetical**, which means what you choose in these questions will not affect your payment in any way.

- We will ask you 20 hypothetical questions about gambles over money. For example, you might be asked to choose between: (a) $50 for sure, or (b) a 50% chance of $110 and a 50% chance of $0.
- We will ask you 20 hypothetical questions about gambles over lottery tickets (where there are 1,000 lottery tickets, one of which earns a prize of $1,000). For example, you might be asked to choose between: (a) 50 tickets for sure, or (b) a 50% chance of 110 tickets and a 50% chance of 0 tickets.

You must answer all 40 hypothetical questions to complete part 2 of the study and earn a part 2 bonus payment.

Figure E.11: Understanding Questions for Instructions, Screen 1

You must answer this question correctly in order to remain in the study. Which of the following statements describes the bonus payments in this study?

- In this study, I will earn $2.50 for sure in part 1. I will not have the opportunity to earn any bonus payments.
- In this study, I will earn $2.50 for sure in part 1. I may also earn a "part 1 bonus" payment. I will earn a "part 2 bonus" payment if I complete part 2 of the study.
- In this study, I will earn $2.50 for sure in part 1. I may also earn a "part 1 bonus" payment. I will also earn $2 for sure in part 2.

You must answer this question correctly in order to remain in the study. What will you do in part 2 of the study?

- I will answer an unknown number of questions about an unknown topic.
- I will answer **40 hypothetical questions** about gambles, and what I choose **will not affect** my payment (except that I must answer all 40 questions to earn a "part 2 bonus" payment).
- I will answer **40 questions** about gambles, and what I choose in each question **will affect** my "part 2 bonus" payment.

>>

Figure E.12: Instructions, Screen 2

As was noted on the last screen, you cannot complete part 2 of the study now. You can only complete part 2 at some point in the future.

In particular, at the end of part 1 of the study, you will receive a link to complete part 2 of the study. The link will become active at 12:01am Eastern Time (ET) on some day in the future and remain active until 11:59pm Eastern Time (ET) a week later. The day that the link will become active will be randomly determined for you during this part of the study. **You are required to complete part 2 of the study in one sitting within that week to earn your part 2 bonus payment.** That is, you must click the link during the week that it is active and complete the survey within one hour of initially clicking the link to earn your part 2 bonus payment.

In this part of the study, we will ask you questions about how you value receiving a set of three reminder emails to complete part 2 of the study during the week that it is available for you to complete. **Based on random chance and your choices in this part of the study, you may receive a set of three reminder emails through the MTurk platform.**

If you do not receive this set of three reminder emails, we will not send you any reminders to complete part 2 of this study.

Figure E.13: Understanding Questions for Instructions, Screen 2

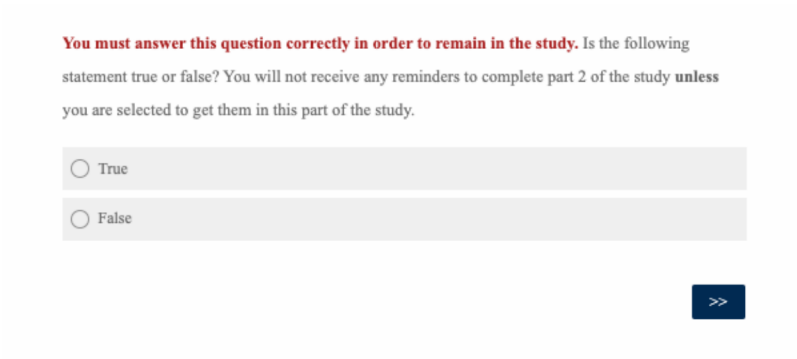


Figure E.14: Instructions, Screen 3

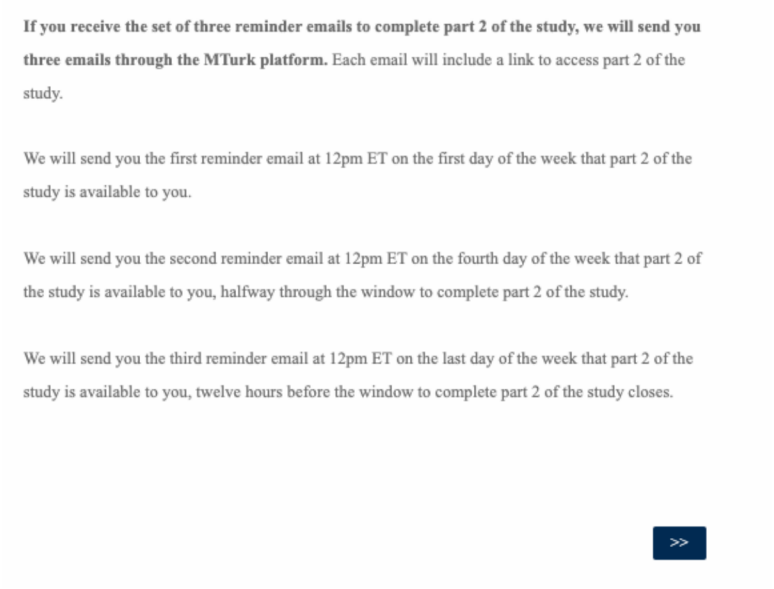


Figure E.15: Instructions, Screen 4

We will now ask you to make decisions about how you value getting the set of three reminder emails to complete part 2 of the study.

You will go through 16 screens of decisions about how you value the reminder emails. There are four possible part 2 bonus payments that you may be eligible to receive and four possible one-week windows in which part 2 of the study may become available to you. (These four possible bonuses and four possible weeks generate 16 possible combinations.) Each possible combination is equally likely to be selected.

Each screen will apply to a different possible part 2 bonus payment for completing part 2 of the study and a different week when part 2 of the study might be available to you.

Because they are randomly selected, you cannot affect which part 2 bonus payment you are eligible to receive or which week part 2 of the study will be available to you. Instead, you will make decisions on the following screens assuming that a given part 2 bonus payment and a given week have been randomly selected. Your responses for a given part 2 bonus payment and week only apply if that part 2 bonus payment and that week are randomly selected.

What this means is that it is in your best interest to make each decision carefully and honestly.

In addition to the decisions you make on the following screens, there is a chance that the computer will randomly assign you to either get the set of three reminder emails or not get the set of three reminder emails.

Any part 1 bonus payment you may earn today and any part 2 bonus payment you may earn from completing part 2 of the study will be sent to you three days after the week in which you may complete part 2 of the study ends.

>>

Figure E.16: Instructions, Screen 5

There will be 16 screens with decisions. On each screen, there will be a table of 33 rows. Each of the rows is its own separate decision. Any one of the decisions could be the one that is selected to determine your outcomes from this study. If one of the decisions is selected, then:

- If you choose the option on the left, you will NOT get the three reminder emails. You will also receive the part 1 bonus payment listed in the option on the left.
- If you choose the option on the right, you will get the three reminder emails. You will also receive the part 1 bonus payment listed in the option on the right.

On each of the 16 screens, you will be asked to indicate which option you prefer in each decision. To speed things up, we have made it so that you can simply click on the decision where you want to switch from choosing the option on the left to the option on the right. What you choose for each decision will be highlighted in orange.

Figure E.17: Understanding Questions for Instructions, Screen 5

You must answer this question correctly in order to remain in the study. If one of the decisions is selected and you chose the option on the LEFT in that decision, what would happen?

- ○ You will NOT get three reminder emails. You will get the amount of money listed in the option on the left.
- ○ You will get three reminder emails. You will get the amount of money listed in the option on the right.

You must answer this question correctly in order to remain in the study. If one of the decisions is selected and you chose the option on the RIGHT in that decision, what would happen?

- ○ You will NOT get three reminder emails. You will get the amount of money listed in the option on the left.
- ○ You will get three reminder emails. You will get the amount of money listed in the option on the right.

>>

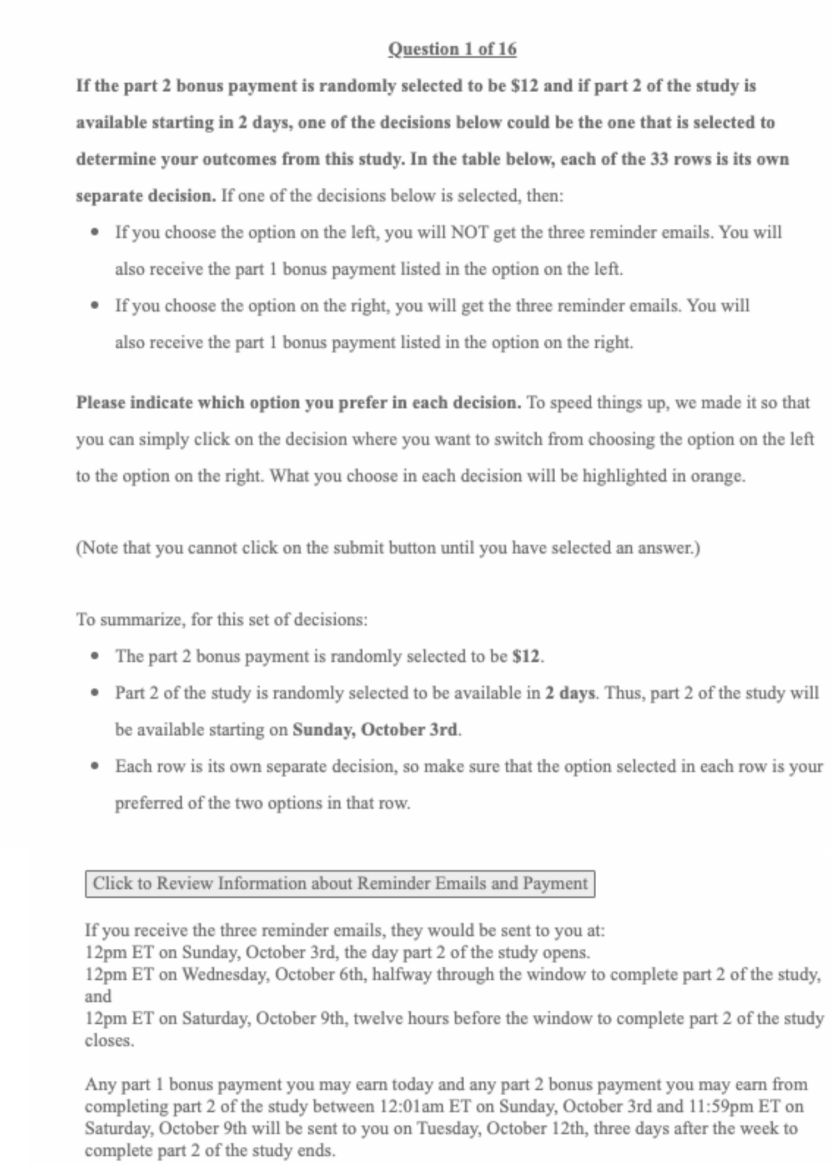
Figure E.18: Multiple Price List Attention Check

Before you make any decisions, we have one more attention check. To signal to us that you are reading all instructions, simply click the continue button below, without clicking on anything in any of the 33 rows below. If you click on any of the rows below, you will be excluded from the study. The decisions will start on the screen after this one.

	NO REMINDERS +		GET REMINDERS +
DECISION 1:	$12.00	OR	$0.00
DECISION 2:	$11.25	OR	$0.00
DECISION 3:	$10.50	OR	$0.00
DECISION 4:	$9.75	OR	$0.00
DECISION 5:	$9.00	OR	$0.00
DECISION 6:	$8.25	OR	$0.00
DECISION 7:	$7.50	OR	$0.00
DECISION 8:	$6.75	OR	$0.00
DECISION 9:	$6.00	OR	$0.00
DECISION 10:	$5.25	OR	$0.00
DECISION 11:	$4.50	OR	$0.00
DECISION 12:	$3.75	OR	$0.00
DECISION 13:	$3.00	OR	$0.00
DECISION 14:	$2.25	OR	$0.00
DECISION 15:	$1.50	OR	$0.00
DECISION 16:	$0.75	OR	$0.00
DECISION 17:	$0.00	OR	$0.00
DECISION 18:	$0.00	OR	$0.75
DECISION 19:	$0.00	OR	$1.50
DECISION 20:	$0.00	OR	$2.25
DECISION 21:	$0.00	OR	$3.00
DECISION 22:	$0.00	OR	$3.75
DECISION 23:	$0.00	OR	$4.50
DECISION 24:	$0.00	OR	$5.25
DECISION 25:	$0.00	OR	$6.00
DECISION 26:	$0.00	OR	$6.75
DECISION 27:	$0.00	OR	$7.50
DECISION 28:	$0.00	OR	$8.25
DECISION 29:	$0.00	OR	$9.00
DECISION 30:	$0.00	OR	$9.75
DECISION 31:	$0.00	OR	$10.50
DECISION 32:	$0.00	OR	$11.25
DECISION 33:	$0.00	OR	$12.00

>>

Figure E.19: Example Multiple Price List Instructions with Incentive Level of $12 and Delay of 2 Days



Question 1 of 16

If the part 2 bonus payment is randomly selected to be $12 and if part 2 of the study is available starting in 2 days, one of the decisions below could be the one that is selected to determine your outcomes from this study. In the table below, each of the 33 rows is its own separate decision. If one of the decisions below is selected, then:

- If you choose the option on the left, you will NOT get the three reminder emails. You will also receive the part 1 bonus payment listed in the option on the left.
- If you choose the option on the right, you will get the three reminder emails. You will also receive the part 1 bonus payment listed in the option on the right.

Please indicate which option you prefer in each decision. To speed things up, we made it so that you can simply click on the decision where you want to switch from choosing the option on the left to the option on the right. What you choose in each decision will be highlighted in orange.

(Note that you cannot click on the submit button until you have selected an answer.)

To summarize, for this set of decisions:

- The part 2 bonus payment is randomly selected to be **$12**.
- Part 2 of the study is randomly selected to be available in **2 days**. Thus, part 2 of the study will be available starting on **Sunday, October 3rd**.
- Each row is its own separate decision, so make sure that the option selected in each row is your preferred of the two options in that row.

Click to Review Information about Reminder Emails and Payment

If you receive the three reminder emails, they would be sent to you at:
12pm ET on Sunday, October 3rd, the day part 2 of the study opens.
12pm ET on Wednesday, October 6th, halfway through the window to complete part 2 of the study, and
12pm ET on Saturday, October 9th, twelve hours before the window to complete part 2 of the study closes.

Any part 1 bonus payment you may earn today and any part 2 bonus payment you may earn from completing part 2 of the study between 12:01am ET on Sunday, October 3rd and 11:59pm ET on Saturday, October 9th will be sent to you on Tuesday, October 12th, three days after the week to complete part 2 of the study ends.

This figure shows the text that appears on the screen of a multiple price list decision. (The next figure shows the actual multiple price list that participants faced.) Participants saw a version of these screen 16 times for every combination of the four possible incentives and four possible delays. The order of these 12 MPLs was randomized at the participant level. In addition to the instruction text, the page has a clickable button between the instructions and the multiple price list which summarizes information about the reminder emails and payment; the figure above shows what it looks like when the button has been pressed.

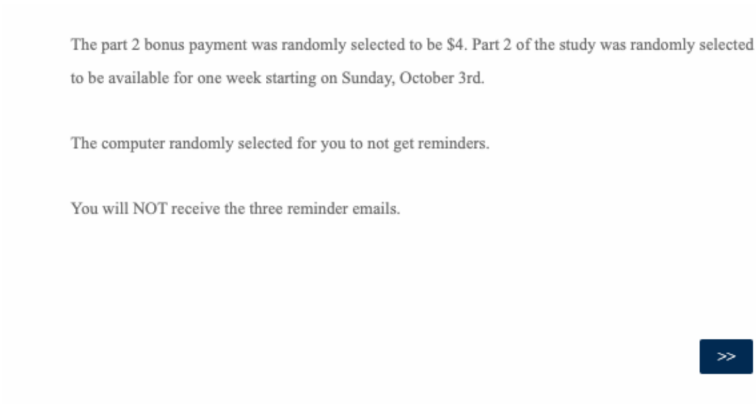
Figure E.20: Example Multiple Price List with Incentive Level of $12 and Delay of 2 Days

	NO REMINDERS +		GET REMINDERS +
DECISION 1:	$12.00	OR	$0.00
DECISION 2:	$11.25	OR	$0.00
DECISION 3:	$10.50	OR	$0.00
DECISION 4:	$9.75	OR	$0.00
DECISION 5:	$9.00	OR	$0.00
DECISION 6:	$8.25	OR	$0.00
DECISION 7:	$7.50	OR	$0.00
DECISION 8:	$6.75	OR	$0.00
DECISION 9:	$6.00	OR	$0.00
DECISION 10:	$5.25	OR	$0.00
DECISION 11:	$4.50	OR	$0.00
DECISION 12:	$3.75	OR	$0.00
DECISION 13:	$3.00	OR	$0.00
DECISION 14:	$2.25	OR	$0.00
DECISION 15:	$1.50	OR	$0.00
DECISION 16:	$0.75	OR	$0.00
DECISION 17:	$0.00	OR	$0.00
DECISION 18:	$0.00	OR	$0.75
DECISION 19:	$0.00	OR	$1.50
DECISION 20:	$0.00	OR	$2.25
DECISION 21:	$0.00	OR	$3.00
DECISION 22:	$0.00	OR	$3.75
DECISION 23:	$0.00	OR	$4.50
DECISION 24:	$0.00	OR	$5.25
DECISION 25:	$0.00	OR	$6.00
DECISION 26:	$0.00	OR	$6.75
DECISION 27:	$0.00	OR	$7.50
DECISION 28:	$0.00	OR	$8.25
DECISION 29:	$0.00	OR	$9.00
DECISION 30:	$0.00	OR	$9.75
DECISION 31:	$0.00	OR	$10.50
DECISION 32:	$0.00	OR	$11.25
DECISION 33:	$0.00	OR	$12.00

>>

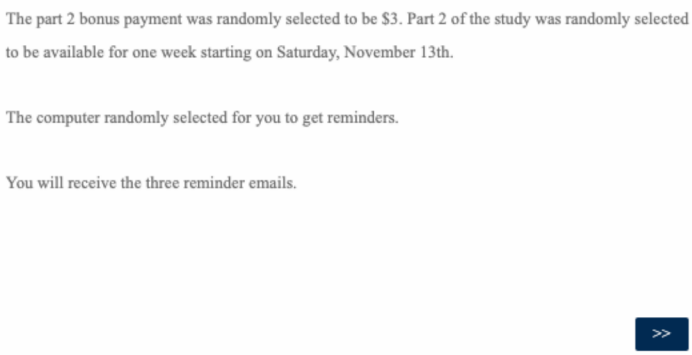
This figure shows the multiple price list that participants saw on the same screen as the text in the previous figure. This is the high incentive treatment; the other version of this multiple price list has a maximum bonus of $4 and increments by $0.25, instead of a maximum bonus of $12 and increments of $0.75 as above.

Figure E.21: Part 2 Information: No reminder emails



Participants who would not receive reminder emails were shown a version of this screen after completing the 16 multiple price list screens; the part 2 bonus and availability was randomly selected and varies across participants.

Figure E.22: Part 2 Information: Reminder emails



Participants who would receive reminder emails were shown a version of this screen after completing the 16 multiple price list screens; the part 2 bonus and availability was randomly selected and varies across participants.

Figure E.23: Link Screen: No reminder emails

The following link will take you to part 2 of the study and be active from 12:01am ET on Sunday, October 3rd to 11:59pm ET on Saturday, October 9th.

https://wharton.qualtrics.com/jfe/form/SV_7QgPF0gZN4ePx8q

If you complete part 2 of the study during this week, you will receive an additional part 2 bonus payment of $4.

Please make a note of this link. You will not receive any reminders to complete part 2 of the study.

Participants who would not receive reminder emails were shown this screen at the end of part 1 of the study.

Figure E.24: Link Screen: Reminder emails

The following link will take you to part 2 of the study and be active from 12:01am ET on Saturday, November 13th to 11:59pm ET on Friday, November 19th.

https://wharton.qualtrics.com/jfe/form/SV_cwg2HYah4aDCOMK

If you complete part 2 of the study during this week, you will receive a part 2 bonus payment of $3.

Please make a note of this link. You will also be sent this link when you are sent reminder emails to complete part 2 of the study.

Participants who would receive reminder emails were shown this screen at the end of part 1 of the study.

Figure E.25: Demographic Information

To complete part 1 of this study, please answer the short follow-up survey below. None of your answers on this follow-up survey will influence your payments in any way.

***How old are you (in years)?**

***What is your gender?**

- Male
- Female
- Other
- Prefer not to say

***Please select the highest level of education that you have completed.**

- Elementary School
- Middle School
- High School or equivalent
- Some college
- College Graduate with Associate's Degree (2 year)
- College Graduate with Bachelor's Degree (4 year)
- Master's Degree (MS)
- Doctoral Degree (PhD)
- Professional Degree (MD, JD, etc.)
- Other

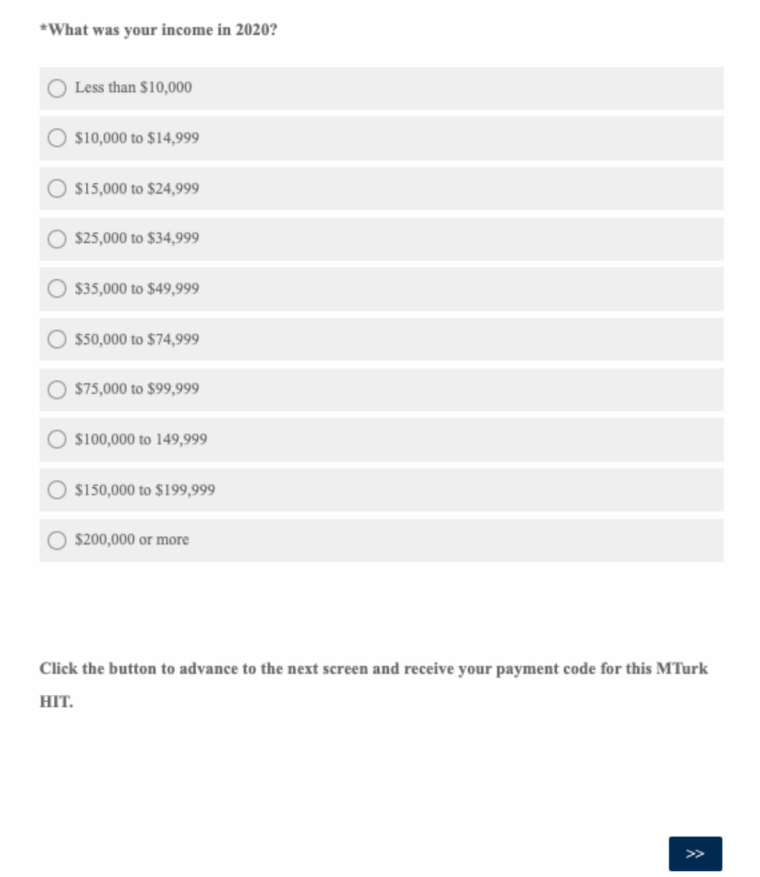
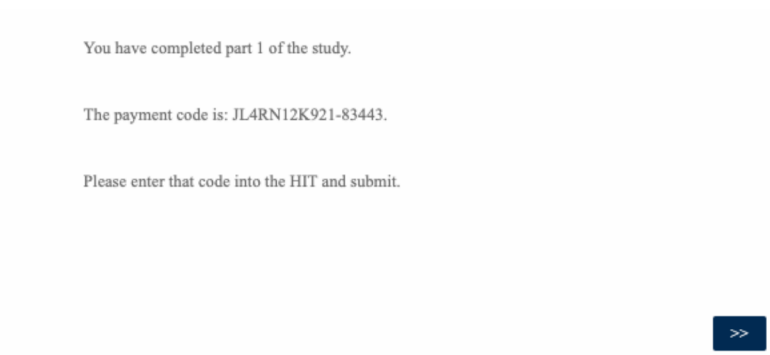


Figure E.26: Final Screen



Participants saw their MTurk payment code as their final screen before exiting the study.