GATE LYON SAINT-ÉTIENNE
UMR 5824

93, chemin des Mouilles
69130 Ecully - France

Maison de l'Université, Bâtiment B
10, rue Tréfilerie
42023 Saint-Etienne cedex 02 - France

http://www.gate.cnrs.fr
gate@gate.cnrs.fr

WP 2314 – September 2023

You Will not Regret it: On the Practice of Randomized Incentives

Brice Corgnet, Roberto Hernán González

Abstract:

Scholars have long emphasized the importance of setting specific goals and providing clear expectations as a key driver of work performance. Yet, a casual observation of actual compensation practices suggests otherwise. Using both lab and field experiments, we show that randomization of bonus targets and piece rates can lead to higher performance than the best-available deterministic scheme. Furthermore, this effect can be sustained over time. We show that part of this effect is explained by regret motives, which lead workers to exert extra effort to avoid missing out on potential pay.

Keywords:
Randomized incentives, bonuses and piece rates, regret, experiments

JEL codes:
C91, C93, D86, D91, M52

Working Paper

You Will not Regret it:

On the Practice of Randomized Incentives

Brice Corgnet and Roberto Hernán González†

Abstract

Scholars have long emphasized the importance of setting specific goals and providing clear expectations as a key driver of work performance. Yet, a casual observation of actual compensation practices suggests otherwise. Using both lab and field experiments, we show that randomization of bonus targets and piece rates can lead to higher performance than the best-available deterministic scheme. Furthermore, this effect can be sustained over time. We show that part of this effect is explained by regret motives, which lead workers to exert extra effort to avoid missing out on potential pay.

Keywords: Randomized incentives, bonuses and piece rates, regret, experiments

JEL codes: C91, C93, D86, D91, M52

† Brice Corgnet: emlyon business school, GATE UMR 5824, F-69130 Ecully, France. Email: corgnet@em-lyon.com. Roberto, Hernán-González, CEREN EA 7477, Burgundy School of Business, Université Bourgogne Franche-Comté, Dijon, France. Email: Roberto.hernan-gonzalez@bsb-education.com. Brice Corgnet acknowledges that this research was performed within the framework of the LABEX CORTEX (ANR-11-IDEX-007) operated by the French National Research Agency. Roberto Hernán-González also acknowledges financial support from ISITE-UBFC International Coach Fellowship. We are thankful to seminar and workshop participants at Ghent, Alicante, Lyon and IESEG for comments on this manuscript. A first version of these results was presented at Nottingham University in 2016 under the title "In the Wild Wild Lab".

1. Introduction

Business scholars have repeatedly advised practitioners to set specific goals and provide clear expectations in order to enhance work performance (e.g., Nadler and Lawler, 1977; Vroom, 1987; Harter et al., 2003; Kim and Mauborgne, 2003; Rock, 2008; Latham, 2012). A common recommendation to managers is thus to "make the performance-reward link clear and objective" (Rock, 2008; Lam et al., 2022). The study of incentives in economics also emphasizes the importance of avoiding noisy production measures because they increase the cost of motivating workers by creating additional risk (see Holmström, 1979; Milgrom and Roberts, 1992; Laffont and Martimort, 2002; Besanko et al., 2017). Furthermore, the theory of incentives has overwhelmingly focused on deterministic schemes that stipulate the exact rewards paid for a given level of production (see Grossman and Hart, 1983; Fudenberg and Tirole, 1991; Laffont and Martimort, 2002; Bolton and Dewatripont, 2004; Dranove et al., 2017).[1] Although randomized schemes can be optimal in the presence of informational asymmetries and risk-sharing issues (see e.g., Baron and Myerson, 1982; Arnott and Stiglitz, 1988; Strausz, 2006; Herweg et al., 2010; Jehiel, 2015; Kadan et al., 2017; Ederer et al.,, 2018), they have often been dismissed for practical reasons, as expressed in Brito et al., (1995, p. 190): "Third, it is unclear whether randomization is only a theoretical curiosity or whether it could be implemented in practice." Yet, behavioral scholars have called for an empirical examination of randomized schemes (Babcock et al., 2012).

Upon casual observation of actual compensation practices, it appears that randomized incentives may be widespread. For instance, various studies have documented the use of randomized incentive schemes in the gig economy to motivate workers through random gifts (Huws et al., 2017; Abraham et al., 2018; Broughton et al., 2018; Hawkins, 2018). In such cases, workers are uncertain about the specific level of performance required to earn the additional reward. Thus, unlike other labor contracts in which the noisy relationship between performance and rewards might be attributed to the absence of precise and verifiable performance metrics (Baker et al., 1994; Hart and Moore, 1999; Tirole, 1999; Holmström, 2017), surprise rewards deliberately introduce noise into the link between performance and pay. This practice, which appears to contradict common recommendations, could nevertheless be profitable.

However, there is no compelling evidence that randomized incentive schemes can be implemented profitably. This lack of evidence may be due to the emphasis placed by economic theorists on

[1] The space dedicated to random incentive schemes in the theory of incentives textbook of Laffont and Martimort (2002) is about 4 pages (that is less than 1% of the material covered) and the same focus on deterministic contracts is observed in major textbooks (Fudenberg and Tirole, 1991; Bolton and Dewatripont, 2004). The working paper of Eric Maskin on randomized incentive schemes (1981) has never been published. After inquiry, we could not obtain a copy from the author.

deterministic schemes, coupled with the widespread recommendation among business scholars to avoid randomization. In practice, randomized schemes present several challenges, as workers may oppose them due to the additional level of risk they introduce and the potential lack of trust in the randomization procedures used by their employers.

In this paper, we investigate whether randomized incentive schemes can be implemented profitably. To test this, we conducted a series of lab and field experiments that manipulated the randomness of bonus and piece rates incentives. Specifically, we examined a setting with a single work task and a single incentive contract, involving either bonuses or piece rates. The incentive literature predicts that randomized incentives are of no value in this context because they cannot alleviate informational asymmetries and risk-sharing issues (EUT-Hypothesis).

We test our hypothesis using a between-subject design in which workers undertake an effortful summation task for which they are paid either a bonus (Bonus treatments) or a piece rate (Piece Rate treatments). In the bonus treatments, a monetary reward is paid if a performance target is achieved, and nothing is paid otherwise. In the Known-Target treatment, one out of three possible targets (low, medium, and high) is picked at random before starting the task and is revealed to the worker. In the Random-Target treatment, the target is randomly picked from the same set as in the Known-Target treatment but is only revealed to workers after completing the task. In the Known-Piece Rate treatment, one out of three possible piece rates (low, medium, and high) is randomly selected before starting the task, whereas in the Random-Piece Rate treatment, the piece rate is only known after the task is completed.

In the lab experiment, we recruited 274 participants from a major US lab who were randomly assigned to the Bonus treatments. Average task production on a 30-minute summation task was 30.5% higher in Random-Target than in Known-Target, and on-the-job leisure (measured by the time spent on recreational websites during the task) was three times lower. Furthermore, the average level of production in Random-Target was 12.0% higher than that of the most effective target, which was the medium value, in Known-Target. Additionally, the on-the-job-leisure level was half as low.

In the field experiment, which was conducted in a major online gig-work platform, we aimed at replicating and extending the previous findings. We recruited 648 workers for a six-week experiment. In Week 0, workers were informed about the duration of the study and completed a series of tasks. In Week 1, workers were randomly assigned to one of the four treatments, and remained in the same incentive treatment for four consecutive weeks: Known-Target, Random-Target, Known-Piece Rate and Random-Piece Rate. In the final week, two thirds of the workers were randomly assigned to one of two Choice treatments. In the Choice-Target (Choice-Piece Rate), workers could choose whether to complete the task under Known-Target or Random-Target (Known-Piece Rate or Random-Piece Rate). In this longitudinal field setting where workers could quit the task or not show up in a given week, we replicated the lab experiments results. Average task production on the 10-minute summation task was 16.3% higher in Random-Target than in Known-Target. Furthermore, Random-Target achieved an average level of production that was 12.6% higher than the most effective deterministic target, which was again the medium value. Extending the lab results to the case of piece rates, we found that average production was 12.5% higher in Random-Target than in Known-Target, and 4.5% higher than in the most effective deterministic piece rate, which was the medium value. Furthermore, these results were sustained over time. Similar to the lab experiment, we did not find treatment differences in levels of work satisfaction. Furthermore, we did not find treatment differences in attrition levels, indicating that randomized incentives did not impact workers' likelihood of showing up for work.

These results show that randomized incentives consistently improve upon the most profitable deterministic incentive scheme. Although this finding contradicts standard predictions in incentive theory, it is consistent with a model in which workers exhibit regret motives. Specifically, randomized schemes can induce regretful workers to exert a higher level of effort because of the fear of missing out on higher pay. In line with this conjecture, we show that the positive impact of randomization can be partly accounted for by regret motives. By eliciting regret using various scales, we show that it significantly explains the impact of randomization.

Regret can boost profits over a sustained period of time as long as employers do not have to fully compensate workers for its negative emotional impact. In our setup, this is the case because, regardless of whether workers decided to complete the task or not, they received feedback about the realized value of bonus targets and piece rates before the end of the weekly task. Ignoring this feedback would be costly to workers, as they would have to leave the platform and forego the corresponding fixed payment for the task. In cases where workers have access to an alternative task that effectively eliminates regret, we expect them to avoid the randomized incentive scheme. This is the situation captured in the Choice-Target and Choice-Piece Rate treatments, where 83.2% of workers decided to stay away from randomized incentive schemes, independently of what treatment they had been assigned to in the previous four weeks. Therefore, employers benefit the most from the implementation of randomized incentives when they do not offer standard deterministic contracts simultaneously. Thus, we anticipate that randomized schemes will be particularly successful if implemented within organizations. This is because employers can ensure that workers do not have access to regret-free alternatives beyond those encountered during an active and costly job search.[2] That said, our field experiment demonstrates the value of randomized incentives in repeated relationships with contractors who have immediate access to multiple job offers.

2. Related literature

2.1. Theoretical literature

Although incentive theory has focused on deterministic contracts (see Holmström, 1979; Grossman and Hart, 1983; Fudenberg and Tirole, 1991; Laffont and Martimort, 2002; Bolton and Dewatripont, 2004), it has long been recognized that randomized mechanisms can be optimal when adverse selection issues are present (Baron and Myerson, 1982; Arnott and Stiglitz, 1988; Strausz, 2006; Herweg et al., 2010; Kadan et al., 2017). For example, Strausz (2006) considers the case of a principal setting a production target in exchange for a bonus payment, involving two types of agents: one that is risk-neutral and the other that is risk-averse. In that context, offering a contract with a deterministic target and another one with randomized targets can reduce the cost of achieving a separating equilibrium because it will help dissuade the risk-averse workers to pick the randomized contract crafted for the risk-neutral agent. Another strand of the literature has shown that, in the case of a multitasking framework in which the agent has private information on the cost associated with the task, randomized schemes can dominate deterministic schemes because they alleviate gaming issues (e.g., Ederer et al., 2018). Gaming issues occur because risk-averse agents tend to focus on the easier task but randomization forces them to rebalance their effort across tasks, which is beneficial to the principal. Furthermore, Jehiel (2015) has shown that randomized schemes can be optimal when the principal has private information over multiple dimensions, such as the monitoring technology and the difficulty of the task, that impact the effort of the agent. In the management literature, Sillince et al., (2010) emphasize that ambiguity can be a powerful tool for creating negotiation and compromise within organizations. By leaving multiple interpretations open, ambiguity can provide space for organizational members with differing goals to remain motivated. Finally, an earlier theoretical work has shown that randomized contracts could be used as a mechanism to improve risk-sharing when risk tolerance varies with the level of effort exerted by the agent (Gjesdal, 1982).

Our work contributes to the theoretical literature on incentives by demonstrating that in the presence of regret motives, randomized schemes can outperform deterministic schemes in situations where screening, gaming, and risk-sharing issues are absent (Bell, 1981; Loomes and Sugden, 1982; Bleichrodt and Wakker, 2015). One key implication is that randomized schemes are not an intellectual

[2] Yet, the worker's disutility associated with regret should not exceed the cost of searching for a regret-free alternative.

curiosity (Brito et al., 1995) whose effectiveness relies on specific informational and institutional assumptions. Instead, they are widely relevant because their success hinges upon widespread behavioral motives.

2.2. Empirical literature

The focus of incentive theory on deterministic schemes might explain the dearth of applied research on randomized incentives. Furthermore, the general belief in the management literature regarding the optimality of transparent incentive schemes might have dissuaded companies to experiment with randomized schemes. In economics, researchers have focused on studying the impact of exogenous shocks on wages and production. Charness and Levine (2007) and Rubin and Sheremeta (2016) have shown that random shocks tend to reduce work performance in gift-exchange experiments in which a principal pays a fixed wage to an agent who then decides on the level of effort to exert on the task. In Rubin and Sheremeta (2016), this is the case because the production shock makes it more difficult for the principal to evaluate the extent of the gift of workers. This weakens reciprocity motives and gift-exchange. In Charness and Levine (2007), wage shocks weaken gift-exchange because generous principals might end up paying low wages. Yet, agents valued the intentions of the principal and were still willing to reciprocate by exerting high effort when principals picked a high wage that was affected by a negative shock. Brownback and Kuhn (2019) report similar results and show that principals believe lucky agents who receive a positive shock on production will work harder than unlucky agents.

In addition to gift-exchange games that do not consider pay for performance, experimental studies using real-effort tasks have assessed the impact of random production shocks in principal-agent relationships involving piece rates (see e.g., Sloof and Van Praag, 2010), linear contracts (Corgnet and Hernán-González, 2019) and bonuses (DellaVigna and Pope, 2018). Sloof and Van Praag (2010) and Corgnet and Hernán-González (2019) report evidence that random shocks can have a positive effect on work performance, which can be explained, as in Gjesdal (1982), by the risk attitudes of prudent agents.

In the eighteen incentive schemes studied in DellaVigna and Pope (2018), two of them introduce a random component ("risk-aversion and probability weighing schemes"). These lottery treatments reward workers with a \$1 (2¢) additional piece rate that is paid 1% (50%) of the time. The 1%-lottery (50%-lottery) treatment led to a slightly lower level of performance, -6.6% (-2.6%) than the benchmark piece rate rewarding 1¢ for sure. Given that people typically overweight small probabilities, the results on the 1%-lottery treatment stress that probability distortions might not have a substantially impact on effort. This is in line with our findings that regret rather than probability distortions explain the performance of randomized incentive schemes. Furthermore, our findings regarding the negative impact of effort regret on workers' performance under randomized schemes helps us to understand the negative effect of the 2%-lottery treatment. In this treatment, effort regret, which refers to the mental cost associated with exerting more effort than one would have if they had known the incentive scheme in advance, is likely to be substantial because workers receive no piece rate 50% of the time, irrespective of their effort level.

As pointed out by the survey of randomized schemes in DellaVigna and Pope (2018), no economic studies had yet implemented randomized incentives in a labor context. Outside of the workplace, lottery incentives offering a stochastic payoff have been offered to boost savings (Filiz-Ozbay et al., 2015) and educational achievements (Levitt et al., 2016).

Most studies using lottery incentives have considered medical applications and report mixed results. Some studies have shown a positive effect on blood donations (Goette and Stutzer, 2020), medicine intake (see Volpp et al., 2008; Kimmel et al., 2012; Sen et al., 2014; see Chang et al., 2017 for a review) and the completion of risk assessments by employees (Haisley et al., 2012). Yet, other medical studies have reported a negative effect on clinicians' survey responses (Halpern et al., 2011), male circumcision uptake in Kenya (Thirumurthy et al., 2016) and chlamydia testing rates (Dolan and Rudisill, 2014).

The mixed results of medical studies and the null results of Levitt et al., (2016) and DellaVigna and Pope (2018) might corroborate the early observation of Brito et al., (1995) that randomized schemes are indeed an intellectual curiosity not worth pursuing. Yet, more positive results have been achieved when lottery incentives have been used to leverage regret motives. Regret lotteries, that differ from standard lotteries by making lost prizes salient to those who did not buy tickets, have been found to enhance participation (Zeelenberg and Pieters, 2004). However, the impact of regret lotteries might be short-lived as shown in a recent study of Imas et al., (2021).

In marketing studies, scholars have also used randomized incentive schemes to trigger persistence and repetitions in experimental tasks (Shen et al., 2015, 2019) following up on early animal research on variable reinforcement schemes (Skinner, 1951; Davison, 1969; Zeilier, 1972; Boakes, 1977; Collins et al., 1983; Forkman, 1993; Hurly and Orseen, 1999). In a series of studies, Shen et al., (2015) show that people were more likely to complete a task (e.g., drink 1.4 liters of water in 2 minutes) for an uncertain reward (e.g., \$1 or \$2 with equal chances) than for a certain reward of lower expected value (e.g., \$1). In addition, Shen et al., (2019) show that uncertain rewards led people to run longer and spend more time completing mental calculations, which is in line with the recent findings of Corgnet et al., (2020). Shen et al., (2019) emphasize that the driving force of any positive effect of randomized incentives is that it appeals to workers because it triggers curiosity regarding the final resolution of uncertainty (e.g., Loewenstein, 1994; Ruan et al., 2018). We refer to this mechanism as the *curiosity hypothesis*. However, none of these studies test the persistence of the effect of randomized schemes, thus not answering the recent critique of Imas et al. (2021).

Overall, the existing literature in behavioral and health sciences provides mixed evidence regarding the motivational effect of lottery incentives and suggests two alternative mechanisms to explain these effects: regret and curiosity.

Our study asks three novel questions: Can randomized incentives be successfully implemented in a labor context? Can the effect of randomized incentives be sustained over time? What are the mechanisms explaining the motivation impact of randomized schemes: regret or curiosity?

In our study, we show that regret motives rather than curiosity play a critical role in explaining work performance under randomized schemes. Indeed, we consider Curiosity-treatments in which we increase the surprise associated with uncertainty resolution by increasing the entropy of rewards (see Corgnet, et al., 2020). However, workers tend to exhibit less effort under the Curiosity-Piece Rate and the Curiosity-Target treatments than under the Random-Piece Rate and the Random-Target treatments. Furthermore, in contrast with the curiosity hypothesis, workers do not report a more positive experience in the Random treatments.

Our work contributes to the empirical literature because it provides the first evidence that randomized incentive schemes can be implemented successfully both in laboratory and longitudinal field experiments. Our study also demonstrates the critical role of regret motives rather than curiosity in explaining the success of randomized schemes.

3. Hypotheses

3.1. Model

3.1.1. Costs

We consider a risk-neutral principal recruiting workers to complete an effortful task. The cost of effort for the worker is assumed to follow a power function so that: $C_i(e_i) = \frac{k}{1+\gamma_i} e_i^{1+\gamma_i}$, where e_i is the number of units of effort produced by the worker, $k \geq 0$ a scaling parameter and $\gamma_i \geq 0$ is a curvature parameter (see e.g., DellaVigna and Pope 2018 for recent empirical validation of this specification). We consider the case of a clerical task in which units of effort deterministically translate into production units. For the sake of simplicity, we will thus use the term effort throughout to refer to either effort or production. We denote $i \in \{L, M, H\} \coloneqq I$ the type of a worker, where $\gamma_H < \gamma_M < \gamma_L$

so that H-workers, for a given level of production, have a strictly lower marginal cost of production than M- and L-workers. That is, H-, M- and L-workers, are the high-, medium- and low-ability workers. Types are workers' private information, and, for simplicity of exposition, we consider that all types are equally likely in the population of workers.

We consider two of the most popular linear and non-linear individual incentive schemes: bonuses and piece rates (Gerhart, Rynes and Fulmer, 2009; Bryson et al., 2012; Gill, Prowse and Vlassopoulos, 2013; DellaVigna and Pope, 2018; Lazear, 2018; Armstrong, 2019). For each incentive scheme, we consider a deterministic (referred to as 'known') and a randomized version. The known (randomized) scheme is such that the target or piece rate is known before starting (after completing) the task. We consider that the principal offers one incentive scheme at a time so that menus of contracts are not contemplated, thus leaving aside screening issues.

3.1.2. Bonus schemes

The principal rewards workers using a bonus (B) that is paid in full if a target production (τ) is achieved. We assume that principals do not know the ability of the workers so that they cannot set a specific target to each type. In this context, we will compare two kinds of incentive schemes. The known scheme randomly picks a target $\tau_i \in \{\tau_L, \tau_M, \tau_H\} \coloneqq \Gamma$, which is known *before* completing the task. The randomized scheme also uses a target $\tau_i \in \Gamma$, but the exact target value is not known until *after* completing the task. In both schemes, one target value is selected at random in Γ, so that each value is equally likely ex ante. Workers know the distribution of the targets.

We assume that $\tau_L < \tau_M < \tau_H$ where $B_- < C_i(\tau_i) < B$ and $C_i(\tau_j) > B$ for any $\tau_j > \tau_i$. That is, in a known scheme, i-workers will be willing to achieve target τ_i but will not be willing to achieve target τ_j ($> \tau_i$) because it requires too much effort given their type.[3] In addition, we assume the cost of effort of i-workers is greater than a lower bound $B_- \geq \frac{2}{3}B$. This implies that, under the randomized scheme, they will not be willing to achieve target τ_i unless they can ensure they receive the bonus with certainty. The expected revenues of a risk-neutral principal are expressed as follows: $\Pi \coloneqq \frac{1}{3}\sum_{i \in I}\left(qe_i - 1_{e_i \geq \tau}B\right)$, where q is the price for each unit of effort (production).

3.1.3. Piece rates

The principal rewards workers using a piece rate (p) per unit of effort so that a worker's total pay is $p \times e_i$. In this context, we will compare two types of incentive schemes. The known scheme randomly

[3] We are implicitly assuming an outside option equal to zero but our results are unaffected by considering a strictly positive outside option.

pays each worker a piece rate $p_i \in \{p_L, p_M, p_H\} \coloneqq P$, which is known *before* completing the task. The randomized scheme also pays a piece rate $p_i \in P$ but the exact piece rate is not known until *after* completing the task. In both schemes, one piece rate value is selected at random in P, so that each value is equally likely ex ante. Workers know the distribution of the piece rates. The expected revenues of a risk-neutral principal are expressed as follows: $\Pi \coloneqq \frac{1}{3}\sum_{i\in I}(q-p_i)e_i$, where $q > p_i$.

3.2. EUT-Hypothesis

3.2.1. Bonuses

We consider the case in which a worker of type i maximizes the following additively separable utility function: $u_i(e_i) \coloneqq 1_{e_i \geq \tau} B - C_i(e_i)$. Under a known scheme, the expected revenues for the principal are given by: $\Pi_L \coloneqq q\tau_L - B$, $\Pi_M \coloneqq \frac{2}{3}(q\tau_M - B)$ and $\Pi_H \coloneqq \frac{1}{3}(q\tau_H - B)$ for targets τ_L, τ_M and τ_H, respectively. For the sake of exposition and in line with our experimental setup, we focus on the case in which the highest expected revenues are achieved under τ_M, and consider that $\bar{\tau} \coloneqq \frac{1}{3}(\tau_L + \tau_M + \tau_H) = \tau_M$.[4]

Under a randomized scheme, workers of a given type i will never be willing to produce at a higher level than τ_i. This is the case because $C_i(\tau_j) > B$ for any $\tau_j > \tau_i$. Thus, L-workers will not produce, and M-workers will produce τ_L at most, and this is achieved when the following condition $\frac{B}{3} \geq C_M(\tau_L)$ holds. H-workers will decide to achieve target τ_H as long as the following two conditions are met: $\frac{B}{3} \geq C_H(\tau_H) - C_H(\tau_M)$ and $\frac{2B}{3} \geq C_H(\tau_H) - C_H(\tau_L)$. If all these conditions are met, the expected production is equal to $\frac{\tau_{L+}\tau_H}{3} = \frac{2}{3}\tau_M$, which is the production level obtained under the most profitable of the known targets (τ_M). It follows that the randomized scheme cannot achieve a higher level of production than the most-profitable known target.

Yet, the randomized scheme can achieve a higher average level of production than the average level obtained across all three known schemes: $\frac{\tau_L+\frac{2}{3}\tau_M+\frac{1}{3}\tau_H}{3}$, which is always lower than $\frac{2}{3}\tau_M$.

So far, we have considered the case in which workers exhibit a linear utility function abstracting away from risk attitudes. Let us denote B_α the level of bonus such that, under risk-aversion, the randomized scheme implements the risk-neutral most-profitable solution, which is when M- and H-workers achieve τ_L and τ_H whereas L-workers do not produce. The level of bonus B_α will increase in the degree

[4] Similar results apply if the highest expected revenues are achieved under τ_L or τ_H.

of risk-aversion of the workers because M-workers will require a risk premium to compensate for the risk associated with receiving the bonus only one-third of the time. This implies that, for a given level of bonus, there will be a level of risk-aversion above which M-workers will decide not to exert effort. In sum, the conditions for the randomized scheme to do as well as the most-profitable known scheme are less likely to be met under risk-aversion.

As a result, if the workers can choose between a known or a randomized scheme for a given level of bonus, more risk-averse workers will be more likely to opt for the known scheme, in which risk is absent, than the randomized scheme.

3.2.2. Piece rates

We assume a worker of type i maximizes the following utility function: $u_i(e_i) \coloneqq p_i e_i - C_i(e_i)$. The incentive compatibility constraint of the randomized scheme is such that the optimal level of effort is equal to: $\left(\frac{\bar{p}}{k}\right)^{\frac{1}{\gamma_i}}$, where $\bar{p} = \frac{1}{3}(p_L + p_M + p_H)$. If we consider, as is the case in our experimental setup, that $\bar{p} = p_M$, the randomized scheme will lead to the same level of production as the known scheme when the piece rate is p_M. The most favorable case for the randomized scheme is when the revenues of the principal are maximized when the known piece rate is p_M, which equals the expected piece rate in the randomized scheme ($\bar{p} = p_M$). However, even in that case, the randomized scheme will lead to production levels and revenues that are equal, but not higher, than those of the most-profitable known scheme.

In the case of risk-averse workers exhibiting a concave power utility function $u_i(e_i) \coloneqq (p_i e_i)^\alpha - C_i(e_i)$ with $0 < \alpha < 1$, which has been repeatedly validated empirically (Stott, 2006), the incentive-compatibility constraint becomes: $e_{i,\alpha}^* = \left(\frac{\alpha p_i^\alpha}{k}\right)^{\frac{1}{1+\gamma_i-\alpha}}$. Generally, randomizing the piece rate has a negative effect on the effort level of risk-averse workers. To illustrate this point, we consider randomizing the piece rate level p_i so that one of the following piece rates is equally likely to apply after completing the task: $p_i - x$, p_i, and $p_i + x$.[5] Then, the impact of the randomization of the piece rate on workers' effort can be assessed by calculating: $\frac{\partial e_{i,\alpha}^{rand^*}}{\partial x} = \frac{\partial}{\partial x}\left(\frac{\alpha((p_i-x)^\alpha + p_i^\alpha + (p_i+x)^\alpha)}{3k}\right)^{\frac{1}{1+\gamma_i-\alpha}} < 0$. Therefore, risk-averse workers will tend to produce less under randomization. A risk-neutral worker

[5] In this case, the average piece rate paid by the principal is the same whether the piece rate is randomized or not. It follows that any difference in workers' production will determine the impact of randomization on the principal's revenues. This is the case that applies to p_M in our experimental setup.

$(\alpha = 1)$ will produce the same as under the known scheme: $e_{i,0}^{*} = \left(\frac{p_i}{k}\right)^{\frac{1}{1+\gamma_i}}$ and achieve a higher level of utility than risk-averse workers.

It follows that risk-averse workers who can choose between a known or a randomized scheme will generally prefer to opt for the known scheme rather than the randomized scheme.

We summarize our conjectures regarding the effect of randomization in the EUT-hypothesis.

EUT-Hypothesis (Expected Utility Theory)

i) Randomized schemes cannot generate higher expected production than the most-profitable known scheme.

ii) Risk-averse workers will prefer known to randomized schemes.

Our EUT-Hypothesis reflects the standard result in the incentive literature that known schemes are optimal in a broad set of applications (see Holmström, 1979; Fudenberg and Tirole, 1991; Laffont and Martimort, 2002; Bolton and Dewatripont, 2004). We pre-registered an alternative behavioral hypothesis according to which randomized schemes will lead to higher levels of performance than known schemes due to regret motives (see 'Random Targets Longitudinal Experiment' (AsPredicted #67348) at https://aspredicted.org/646ji.pdf). We detail our regret hypothesis next.

3.3. Regret-Hypothesis

We put forth that randomization can be especially motivating when workers exhibit regret motives (Bell, 1981; Loomes and Sugden, 1982; Bleichrodt and Wakker, 2015). Indeed, workers might decide to exert more effort in the randomized scheme to avoid missing out on potential earnings. At the same time, workers might also regret having exerted too much effort when, for example, the realized target is substantially lower than their actual performance. We must thus contemplate two types of regret, which we will refer to as *monetary* and *effort* regret. Monetary regret refers to the negative feeling (mental cost) experienced by workers when they realize they could have earned more had they known the incentive scheme in advance. Effort regret refers to the negative feeling experienced by workers when they realized they could have been better off exerting less effort had they known the incentive scheme in advance. We define regret with respect to the level of utility a worker would have achieved had they known the realized incentive scheme. Given this definition of the reference point, there cannot be rejoice in our analysis as one can never surpass the reference level of utility that is achieved under complete information.

3.3.1. Bonuses

In a randomized scheme, we consider that a regretful worker of type i will maximize the following utility function:

$$
\begin{aligned}
E[u_i^r(e_i)] \coloneqq & \frac{1}{3}\sum_{\tau\in\Gamma}(1_{e_i\geq\tau}B - C_i(e_i)) \\
& -\frac{r_m}{3}\sum_{e_i<\tau,\tau\leq\tau_i}\left(u_i(e_i^*|\tau) - \left(0 - C_i(e_i)\right)\right) && [1_1] \\
& -\frac{r_e}{3}\sum_{e_i\geq\tau,\tau\leq\tau_i}\left(u_i(e_i^*|\tau) - \left(B - C_i(e_i)\right)\right) && [1_2] \\
& -\frac{r_e}{3}\sum_{e_i<\tau,\tau>\tau_i}\left(u_i(e_i^*|\tau) - \left(0 - C_i(e_i)\right)\right) && [1_3]
\end{aligned}
$$

Where $u_i(e_i^*|\tau)$ is the utility of workers of type i when knowing the target is τ, thus exerting the optimal level of effort given the value of the target. The first term in the regret function $[1_1]$ is what we refer to as monetary regret, captured with $r_m > 0$, as it measures the increase in utility workers could have obtained from receiving the bonus payment if they had worked more and cleared a target that was achievable ($\tau \leq \tau_i$). The second term of the regret utility function $[1_2]$ is what we call effort regret, captured with $r_e > 0$, as it measures the increase in utility workers could have obtained if they had worked less to achieve the target, thus suffering lower effort costs. The third term $[1_3]$ is also related to effort regret and measures the increase in utility workers could have obtained had they exerted no effort given that the target was not achievable ($\tau > \tau_i$). Our formulation of regret is consistent with Gill and Prowse (2012), which incorporates regret in a tournament incentive scheme using a real-effort task.[6]

Given that workers of type i will never produce more than τ_i, the only possibility for the randomized scheme to achieve a higher production than the most-profitable known scheme is when both M- and H-workers produce τ_M and τ_H, respectively. We know that, under a randomized scheme, H-workers can produce τ_H because the likelihood of obtaining the bonus, as in the case of the known target τ_H, is one. However, it remains to be seen whether M-workers can produce τ_M in the presence of regret. For M-workers, monetary regret comes from not achieving a target that is achievable, that is τ_L or τ_M (see $[1_1]$). Effort regret arises when a target could have been achieved with less effort $[1_2]$ or when effort has been exerted but no bonus was received because the target was not achievable $[1_3]$. M-workers will prefer to achieve τ_M rather than τ_L if $E[u_M^r(\tau_M)] \geq E[u_M^r(\tau_L)]$, that is:

[6] We are not aware of other works introducing regret to explain work performance.

$$\frac{2B}{3} - C_M(\tau_M) - \frac{r_e}{3}\big(C_M(\tau_M) - C_M(\tau_L)\big) - \frac{r_e}{3}C_M(\tau_M)$$
$$\geq \frac{B}{3} - C_M(\tau_L) - \frac{r_m}{3}(B - (C_M(\tau_M) - C_M(\tau_L))) - \frac{r_e}{3}C_M(\tau_L)$$

$$\Leftrightarrow \frac{B}{3} - (C_M(\tau_M) - C_M(\tau_L)) + \frac{r_m}{3}\Big(B - \big(C_M(\tau_M) - C_M(\tau_L)\big)\Big) - \frac{2r_e}{3}\big(C_M(\tau_M) - C_M(\tau_L)\big) \geq 0$$

By definition, we have that $B > C_M(\tau_M) - C_M(\tau_L)$ so that the stated condition is more likely to be satisfied as monetary regret increases, and less likely to be satisfied as effort regret increases.

Because regret can impose mental costs that are unappealing to workers, it is necessary to check whether the participation constraint holds. To that end, we consider producing nothing as the outside option available to workers. Importantly, we assume, as in our experimental setup, that workers receive feedback about the outcome of the randomization of incentives even when they decide not to produce. As a result, workers will also experiment monetary regret when producing nothing. The condition under which M-workers will prefer to achieve τ_M rather than producing zero, $E[u_M^r(\tau_M)] \geq E[u_M^r(0)]$, can thus be stated as follows:

$$\frac{2B}{3} - C_M(\tau_M) - \frac{r_e}{3}\big(C_M(\tau_M) - C_M(\tau_L)\big) - \frac{r_e}{3}C_M(\tau_M) \geq -\frac{r_m}{3}\big(B - C_M(\tau_M)\big) - \frac{r_m}{3}\big(B - C_M(\tau_L)\big)$$

$$\Leftrightarrow \frac{2B}{3} - C_M(\tau_M) + \frac{r_m}{3}\Big(2B - \big(C_M(\tau_M) + C_M(\tau_L)\big)\Big) - \frac{r_e}{3}\big(2C_M(\tau_M) - C_M(\tau_L)\big) \geq 0$$

This condition is automatically satisfied if effort regret is absent and is more (less) likely to be satisfied as r_m (r_e) increases (decreases).

Not only do we have to ensure M-workers achieve target τ_M, but we also have to ensure that H-workers achieve τ_H. H-workers do not settle for target τ_M as long as $E[u_H^r(\tau_H)] \geq E[u_H^r(\tau_M)]$, that is the case when:

$$B - C_H(\tau_H) - \frac{r_e}{3}\big(C_H(\tau_H) - C_H(\tau_M)\big) - \frac{r_e}{3}\big(C_H(\tau_H) - C_H(\tau_L)\big)$$
$$\geq \frac{2B}{3} - C_H(\tau_M) - \frac{r_m}{3}\Big(B - \big(C_H(\tau_H) - C_H(\tau_M)\big)\Big) - \frac{r_e}{3}\big(C_H(\tau_M) - C_H(\tau_L)\big)$$
$$\Leftrightarrow \frac{B}{3} - \big(C_H(\tau_H) - C_H(\tau_M)\big) + \frac{r_m}{3}\Big(B - \big(C_H(\tau_H) - C_H(\tau_M)\big)\Big) - \frac{2r_e}{3}\big(C_H(\tau_H) - C_H(\tau_M)\big)$$
$$\geq 0$$

As in previous conditions, the left-hand side increases in monetary regret (r_m), and decreases in effort regret (r_e).

Similarly, we want to make sure that H-workers do not to settle for target τ_L, which is the case when $E[u_H^r(\tau_H)] \geq E[u_H^r(\tau_L)]$, that is:

$$
\begin{aligned}
B - C_H(\tau_H) - \frac{r_e}{3}\big(C_H(\tau_H) - C_H(\tau_M)\big) - \frac{r_e}{3}\big(C_H(\tau_H) - C_H(\tau_L)\big) \\
\geq \frac{1}{3}B - C_H(\tau_L) - \frac{r_m}{3}\Big(B - \big(C_H(\tau_H) - C_H(\tau_L)\big)\Big) - \frac{r_m}{3}\Big(B - \big(C_H(\tau_M) - C_H(\tau_L)\big)\Big) \\
\Leftrightarrow \frac{2}{3}B - \big(C_H(\tau_H) - C_H(\tau_L)\big) + \frac{r_m}{3}\Big(2B - \big(C_H(\tau_H) + C_H(\tau_M) - 2C_H(\tau_L)\big)\Big) \\
- \frac{r_e}{3}\big(2C_H(\tau_H) - C_H(\tau_L) - C_H(\tau_M)\big) \geq 0
\end{aligned}
$$

Again, the left-hand side increases in monetary regret (r_m), and decreases in effort regret (r_e). A similar condition is obtained when ensuring H-workers do not produce zero, $E[u_H^r(\tau_H)] \geq E[u_H^r(0)]$:

$$
B - C_H(\tau_H) + \frac{r_m}{3}\Big(3B - \big(C_H(\tau_H) + C_H(\tau_M) + C_H(\tau_L)\big)\Big) - \frac{r_e}{3}\big(2C_H(\tau_H) - C_H(\tau_L) - C_H(\tau_M)\big) \geq 0
$$

As a result, we have shown that for high levels of monetary regret, relative to effort regret, the randomized scheme can lead both M- and H-workers to achieve their target so that expected production is at least equal to $\frac{\tau_M+\tau_H}{3}$, which is greater than the expected production obtained in the most-profitable known scheme ($\frac{2}{3}\tau_M$).

In the presence of regret motives, L-workers can also reach target τ_L if it is better for them than producing nothing, which is the case whenever $E[u_L^r(\tau_L)] \geq E[u_L^r(0)]$:

$$
\frac{B}{3} - C_L(\tau_L) + \frac{r_m}{3}\big(B - C_L(\tau_L)\big) - \frac{2r_e}{3}C_L(\tau_L) \geq 0
$$

Again, the left-hand side increases in monetary regret (r_m), and decreases in effort regret (r_e). Thus, we have shown that for large values of monetary regret relative to effort regret, workers of type i will always achieve τ_i. Under known schemes, this outcome can only be obtained when there is not private information about workers' types so that principals can offer type-specific bonus contracts.

3.3.2. Piece rates

Under a randomized piece rate scheme, monetary (effort) regret arises when the piece rate is high (low) because workers who had known the piece rate value would have earned more (worked less). We consider here whether, in the presence of regret, a randomized piece rate can lead to higher production levels than the most-profitable piece rate, which is assumed to be equal to p_M. This is an interesting case because workers are paid the same expected piece rate in both schemes so that any difference in workers' production also directly translates into a difference in the principal's expected revenues.

If we contemplate the case in which regret has an overall positive effect on production, we must consider that the production of a regretful worker is higher than the level of production that would have been achieved knowing the piece rate was p_M. It follows that if the randomized piece rate turns out to be p_M, a regretful worker will experience effort regret because knowing the piece rate ex ante would have led to a lower effort. We thus consider that a regretful worker of type i maximizes the following utility function:

$$\mathrm{E}[u_i^r(e_i)] \coloneqq p_M e_i - C_i(e_i) - \frac{r_m}{3}\Big(u_i^*(e_i^*|p_H) - \big(p_H e_i - C_i(e_i)\big)\Big) - \frac{r_e}{3}\Big(u_i^*(e_i^*|p_L) - \big(p_L e_i - C_i(e_i)\big)\Big) - \frac{r_e}{3}\Big(u_i^*(e_i^*|p_M) - \big(p_M e_i - C_i(e_i)\big)\Big) \quad [2]$$

The incentive compatibility constraint is such that: $e_i^{r*} = \left(\frac{p_M + \frac{r_m}{3}p_H + \frac{r_e}{3}(p_L + p_M)}{k(1 + \frac{r_m}{3} + \frac{2r_e}{3})}\right)^{\frac{1}{\gamma_i}}$. That is, in the absence of regret ($r_m = r_e = 0$), the worker exerts the same effort as under EUT under the randomized scheme so that $e_i^* = \left(\frac{p_M}{k}\right)^{\frac{1}{\gamma_i}}$. In addition, in the case in which monetary and effort regret are the same ($r_m = r_e$), regret does not impact production, that is $e_i^{r*} = e_i^*$. As the level of monetary regret increases, the level of production of a worker of type i converges to the level achieved when the piece rate is p_H because $\lim_{r_m \to \infty} e_i^{r*} = \left(\frac{p_H}{k}\right)^{\frac{1}{\gamma_i}}$ despite receiving an average piece rate of p_M. As a result, there exists a level of monetary regret above which the level of production obtained under the randomized scheme is greater than what is obtained for a known scheme paying all workers p_M. Furthermore, the randomized scheme can also achieve a higher level of production than that achieved under a known scheme when there is no private information about workers' types, in which case principals can offer type-specific piece rate contracts (p_i).

Thus far, we have assumed that the participation constraint held. This is the case as long as we assume, as we did for bonuses, that workers receive feedback about the outcome of the randomization of incentives even when they decide not to produce on the task. In that case, not producing on the task is associated with large monetary regret costs and the absence of effort costs, thus leading to the following outside option: $-\frac{r_m}{3}\big(u_i^*(e_i^r|p_H)\big)$. It is easy to see by comparing this term to equation [2] that the participation constraint will be satisfied whenever monetary regret is sufficiently large or effort regret is sufficiently low.

3.3.3. The choice between known and randomized schemes

Thus far, we have considered a case in which the implicit outside option was to produce nothing. This is the case whenever changing employer implies substantial costs that would make an alternative offer unappealing. In that context, workers cannot eliminate the regret that is inherently present under randomized schemes because they will always receive feedback about actual realizations of the targets and piece rates values. Yet, if we consider that workers have a no-regret outside option available, they will tend to avoid randomized schemes. Here, we consider the case of a regretful worker who can choose between a randomized and a known scheme. For both the randomized bonus and the randomized piece rate, the expected utility of the regretful worker is less than or equal to the level obtained in the absence of regret, that is: $E[u_i^r(e_i^{r*})] \leq E[u_i(e_i^*)]$. In the case of randomized piece rates, it follows from the fact that that the level of production of regretful workers does not necessarily satisfy the incentive compatibility constraint in the absence of regret. In the case of the randomized bonus, it follows from the fact that the level of production of regretful workers can be lower than the expected bonus, as is the case of M-workers achieving τ_M.

Since known schemes eliminate regret, they result in an expected utility level that is at least as high as that obtained under randomized schemes. Consequently, regretful workers will opt for the known rather than the randomized scheme. We summarize our second hypothesis, derived under the presence of regret, as follows.

Regret-Hypothesis

i) Monetary regret will increase workers' production in randomized schemes and effort regret will have the opposite effect. Regret will not impact workers' production in known schemes.

ii) Randomized schemes can generate higher production levels than the most-profitable known scheme in the presence of monetary regret motives.

iii) Regretful workers will prefer known schemes to randomized schemes.

We will test the Regret-Hypothesis along with alternative behavioral theories based on probability distortion, loss aversion and curiosity motives in Section 5.3.

4. Design

To test our hypotheses, we conducted a lab experiment at a major US lab in March-April 2016, as well as a field experiment with workers from a major online gig-work platform in June-July 2021. The field experiment helped us both to replicate and extend the previous findings obtained in the lab. In particular, the field experiment allowed us to pinpoint the underlying mechanisms by eliciting regret, loss aversion and probability distortions.

4.1. Lab experiment

We developed a framework in which participants can undertake a real-effort task while having access to a real-leisure alternative (browsing the Internet) at any time during the experiment (Corgnet et al., 2015). The laboratory setting allows the experimenter to control for potential confounding factors commonly encountered in the field, such as organizational hierarchies or implicit incentives, thus facilitating the detection of incentive effects.[7]

We considered a real-effort summation task that was particularly long, laborious, and effortful (e.g., Dohmen and Falk, 2011; Eriksson et al., 2009; Niederle and Vesterlund, 2007). Participants had to sum up matrices of 36 numbers comprised between 0 and 3 for one hour, during a 30-minute period. After completing a table, participants were immediately informed whether their answer was correct or not. Participants were not allowed to use a pen, scratch paper or calculator. This rule amplified the level of effort participants had to exert to complete the tables correctly.

At any point during the experiment, participants could switch from the work task to the leisure activity that consisted of browsing the Internet. Each activity was undertaken separately, on a different screen so that participants could not sum tables while being on the Internet. Participants were informed that their use of the Internet was strictly confidential.[8] The Internet browser was embedded in the software so that the experimenter could keep record of the exact amount of time participants spent on each activity.[9]

At the beginning of the experiment, participants were informed that they will receive a fixed pay of $6 and a $6 bonus if they were able to complete correctly a certain number of tables, i.e., the target. All participants were informed that their personal target will be randomly selected between three possible values: 13 (low target), 26 (medium), or 39 (high), all with equal chances. To ensure the credibility of the target randomization procedure, a monitor rolled a six-sided die in front of all the participants. Prior to rolling the die, participants indicated on their screen which targets (13, 26, and 39) they wanted to associate with the following pairs of outcomes of the six-sided die roll: (1,2), (3,4) and (5,6).[10]

[7] In our lab setting, workers only face explicit incentives.
[8] Participants were expected to follow the norms set by the university regarding the use of Internet on campus.
[9] The lab policy is to forbid cell phone use inside the lab. This ensures that embedded internet browsing is an accurate measure of on-the-job leisure.
[10] This implies that the experimenter cannot effectively use an unfair die to minimize participants' payments.

Target values were selected based on a previous calibration experiment in which 31 participants were recruited to sum tables for 30 minutes and received a piece rate of 16¢ for each correct answer.[11] The median production in these experiments was 26 tables (see Figure A.1 in Appendix A), while 87.1% of the participants were able to complete 13 tables and only 19.4% of the participants completed 39 tables or more.

We conducted two treatments. In the Known-Target treatment, a target was randomly selected by the computer at the beginning of the experiment and participants were informed about it before starting the 30-minute work task. This target was displayed on the screen along with the number of correct tables completed, so it was clear to participants whether they had reached their target. The Random-Target treatment was identical to the Known-Target treatment except that the Target was selected at the end of the experiment, so participants worked on the task without knowing the exact target that would apply.

The experiment was computerized (using a custom Java software; see Corgnet et al., 2015, for more details) and there was no interaction among participants. The instructions were displayed on the participants' computer screens (see Supplementary Material, Lab Study). Participants had exactly 20 minutes to read the instructions. A 20-minute timer was shown on the laboratory screen. Three minutes before the end of the instructions period, a monitor announced the time remaining and handed out a printed copy of the summary of the instructions. None of the participants asked for extra time to read the instructions. At the end of the 20-minute instruction round, the instructions file was closed, and the experiment started. The interaction between the experimenter and the participants was negligible.

At the beginning of the experiment and before having access to the instructions for the work task, participants could add five one-digit numbers for two minutes. Each correct answer provided 6¢. We used their performance on this task as a measure of their mathematical ability. At the end of the experiment and before payments were made, we asked participants to add again five one-digit numbers for two minutes, but this time they received no monetary rewards. Performance on this task was used as a proxy of participants' intrinsic motivation on the work task. We also obtained other measures of general intrinsic motivation and enjoyment on the task using the Self-Determination Inventory (Ryan, 1982; Ryan, Koestner and Deci, 1991) and the Self-Assessment Manikin (Bradley and Lang, 1994). We also had information about participants' cognitive abilities (Raven's Progressive Matrices test, Raven, 1936; Cognitive Reflection Test, CRT henceforth, Frederick, 2005), risk attitudes (Holt and

[11] The choice of the calibration piece rate was such that the average earning would be similar to the $6 bonus. In the case of a 16¢ piece rate the average earning was $4.4. Similar calibration results were obtained using an additional experiment with a 23¢ piece rate in which case the average earning was $6.8 ($n$ = 32).

Laury, 2022) and personality traits (Big Five, John et al., 1991, 2008) from a survey conducted at the beginning of the school year, that was taken by the majority of the participants in the subject pool of the lab (see Corgnet et al., 2018, for details).

Participants received their earnings in cash at the end of the experiment. Participants earned on average \$11.3 including a \$7 show-up fee. Experimental sessions lasted one hour on average. The full set of instructions is available in the Supplementary Material (Lab Study).

4.2. Field experiment

We recruited 998 workers from Amazon Mechanical Turk for an initial 45-minute task (Week 0). In this week, workers were informed that, if they completed the initial task, they would be invited to five other 30-minute sessions taking place in subsequent weeks (Weeks 1 to 5). However, earnings for each session would be calculated and paid separately, so workers could attend all sessions or just a few. All observations were collected using custom Javascript software programmed by the authors and deployed via Qualtrics. This study was advertised as being part of a research conducted by a major research center. The structure and content of the experimental sessions were pre-registered (https://aspredicted.org/646ji.pdf). Our target was to have 700 workers in Weeks 1 to 4 and we got 648.[12]

In Week 0, workers completed a battery of cognitive tests (e.g., Cognitive Reflection Test, Stroop, Raven, Numeracy), personality tests (e.g., Big Five, General Self-Efficacy, Gambling Related Cognitions), tasks measuring distributive social preferences, math ability, self-reported risk attitudes, and socio-demographics. All these measures are commonly-used in economic experiments (see Supplementary Material, Field Study-Additional Instructions). The average earnings for Week 0 was \$7.6 for tasks that were completed in about 50 minutes.[13] In Week 0, workers were randomly assigned to incentive treatments. However, we never informed them about the treatment they had been assigned to or the tasks they were going to perform in the following weeks to avoid attrition driven by treatment conditions.

In Weeks 1 to 4, workers had to complete the main work task, consisting in adding five one-digit numbers for 10 minutes. Following the work of Gill and Prowse (2012), we use a real-effort task for various reasons. First, this allowed us to study randomized schemes in a work setting in which both monetary and effort regret are relevant, thus extending previous empirical works focusing on lottery

[12] Given the longitudinal nature of our design, we decided not to collect the 7% missing data in the Fall or later that year because of the risk of sudden changes in the COVID situation that could have attracted a very different subject pool.

[13] We report the median recorded completion time to limit the impact of extreme values. Yet, this recorded time likely overestimates the actual time MTurkers spent on the task as they could take breaks during the completion of the experiment.

choices (see e.g., Loomes and Sugden, 1987; Loomes, 1988; Starmer and Sugden, 1989, Starmer, 1992; Starmer and Sugden, 1993; Somasundaram and Diecidue, 2017). Second, using real-effort tasks is likely to exacerbate regret motives (see Van Dijk et al., 1999). In contrast to the laboratory study, we did not include any leisure activities, as our participants were working online and could therefore undertake any other activity they deemed appropriate. After the work task, workers completed some questions adapted from the Intrinsic Motivation Inventory (Gagne and Deci, 1992) as well as a short two-minute addition task with a different piece rate per session. At the end of each week, workers had to complete additional survey questions and tests. The average earnings for Weeks 1 to 4 was \$2.5 for tasks that were completed in about 25 minutes.[14]

Workers were assigned to either Target or Piece Rate treatments. In the Target treatments, a \$3 bonus was paid to all the workers who correctly solved a certain number of sum problems. In the Piece Rate treatments, workers received a piece rate for each correctly solved sum problem. For both types of treatments, we conducted three conditions: Known, Random, and Curiosity. We thus implemented a 2×3 between-subject design. Workers were assigned to the same treatment for Weeks 1 to 4. For each of the six incentive scheme treatments, workers were also randomized across four conditions that varied their weekly schedule for completing the task. In the Preset-Low-Flex treatment, workers had to complete the task on a specific day of the week (Monday, Tuesday, Wednesday, Thursday or Friday) selected at random by the computer. The Low-Flex treatment was the same as Preset-Low-Flex except that workers chose their preferred delivery day in Week 1. In the High-Flex treatment, workers could complete the task any day of a given week (from Monday to Friday). Finally, in the Choice-Flex condition, workers chose in Week 1 whether to participate in the High-Flex condition or the Low-Flex condition. In our analyses, we control for scheduling fixed effects when assessing the impact of incentive treatments.

In the Known and Random treatments, the incentive scheme was randomly selected from a list of three possible values: {50, 80 or 110} for targets and {1.25¢, 3.75¢ or 6.25¢} for piece rates. In the Known treatments, workers learned the randomly selected value before completing the 10-minute work task, whereas in the Random treatments, workers were not informed of the selected value until they had completed the work task. The Curiosity treatments were identical to the Random treatments, except that the value of the target or the piece rate was randomly selected between the lowest and highest values (i.e., any integer between 50 and 110 for targets [61 possible values], and any two-decimal number between 1.25 and 6.25 for piece rates [501 possible values]). In all treatments, workers knew

[14] We report the median recorded completion time to limit the impact of extreme values. Yet, this recorded time likely overestimates the actual time MTurkers spent on the task as they could take breaks during the completion of the experiment.

the set of possible values associated with an incentive scheme, and they knew that each week a value was drawn at random for each worker.

The targets and piece rates were calibrated using results from previous 2-minute versions of the same task performed by the authors on the same online platform. As for the lab targets, we calibrated the medium target value so that the median worker could achieve it, and low and high targets so that about 80% (20%) would be able to achieve it. Piece rates were chosen so as to induce similar earnings to those in the target treatments. Thus, workers completing 80 sums problems would earn \$3 whether they were assigned a medium target or a medium piece rate.

In the last session (Week 5), workers were randomly assigned to three possible conditions, regardless of the treatment they had been assigned to in Weeks 1 to 4. Approximately two thirds of the workers (234 out of 360) had the possibility to choose between knowing the incentive scheme (Choice-Target, 122 workers, or Choice-Piece Rate, 112) ex ante, as in the Known treatments or ex post, as in the Random treatments. The rest of the workers (126 out of 360) were assigned to a task eliciting an incentivized measure of effort regret. In this task, we gave workers multiple price lists to elicit the extent to which they preferred a randomized versus a deterministic bonus scheme for completing the 10-minute adding task. In each list, Option A was a randomized bonus scheme paying a \$3 bonus with 70% chance if a given target was achieved and nothing otherwise, and Option B was a deterministic bonus ranging from \$0.5 to \$3 in increments of 50¢, which was paid whenever the given target was achieved. Workers' choices between the two options thus revealed the *deterministic bonus equivalent* of the randomized bonus scheme. Workers had to complete three multiple price lists that only differed in the target: 1, 40 or 80. To incentivize the task, we randomly selected one of the targets along with one of the six choices in the list and implemented the corresponding decision. Thus, workers had 10 minutes to achieve the selected target number and were paid according to their selected bonus scheme choice. The average earnings for Week 5 was \$2.9 for tasks that were completed in about 20 minutes.[15] The full set of instructions is available in the Supplementary Material (Field Study).

We included several attention checks in the field experiment to increase the reliability of the data (Oppenheimer et al., 2009). Attention checks were embedded in the questionnaires, with a similar format to the other questions, requiring participant to select some specific option (e.g., "Select 'strongly agree' for this item"). Of the 13 attention checks, 3 in Week 0 and 2 in the other weeks, only 4.1% of them were not completed correctly (values ranged from 3.0% in Week 2 to 5.2% in Week 4). This percentage is similar to other MTurk studies and lower than values generally reported for lab

[15] We report the median recorded completion time to limit the impact of extreme values. Yet, this recorded time likely overestimates the actual time MTurkers spent on the task as they could take breaks during the completion of the experiment.

studies (Klein et al., 2014; Hauser & Schwarz, 2016). As preregistered, participants who failed any of the attention checks within a week were excluded from the corresponding analysis.

5. Results

5.1. On the performance of randomized schemes

5.1.1. Lab data

We observe that the Random-Target treatment led to higher production levels (27.16 tables) and lower levels of on-the-job leisure (measured as the percentage of time spent on the Internet, 7.3%) than in all the Known-Target treatments (20.81 tables and 21.7% Internet use; i.e., +30.5% and -66.4% respectively) (see Figure 1). We confirm that these differences are statistically significant in Table 1, where we report linear regressions for standardized levels of production and internet use. We observe that the coefficient test for 'Random-Target' is positive and significant for production (Regression [1], p-value < 0.001) while the coefficient for internet use is negative and also significant (Regression [3], p-value < 0.001).

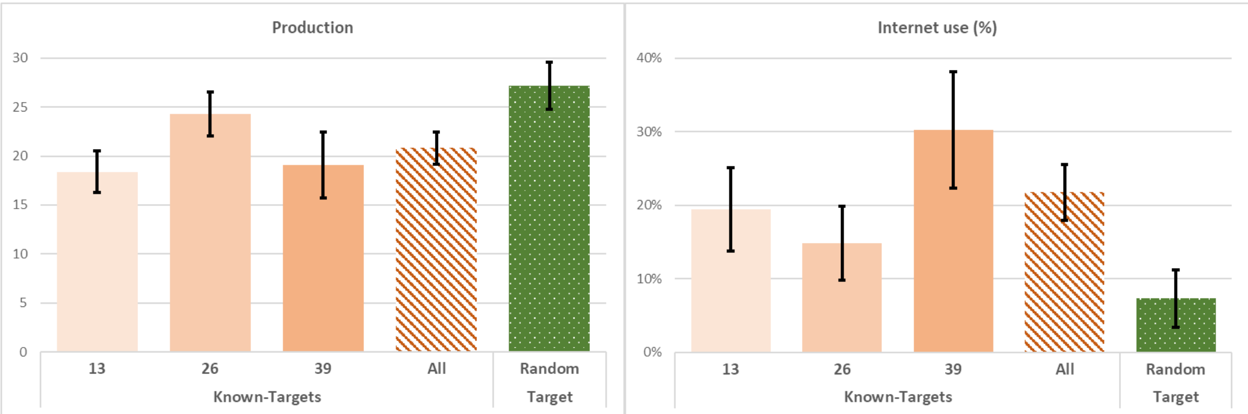


Figure 1. Average production levels (left panel) and internet use (in % of available time for the task) (right panel), with 95% confidence intervals.

Furthermore, the Random-Target treatment led to a higher level of production than any of the known targets (see Figure 1). It outperformed the low and high targets (18.29 and 19.07 tables; p-values <0.001, coefficient tests for 'Random-Target = Known-Target Low' and 'Random-Target = Known-Target High' in Regression [2]). It also increased production by 12.0% with respect to the best-performing known target, which corresponds to the medium target value (p-value < 0.001, coefficient test for 'Random-Target' in Regression [2]). Similar results are obtained when comparing treatments on Internet use (see regressions [3] and [4]). In particular, the Random-Target treatment decreased Internet use by 50.8% compared to the best-performing known target (14.8% internet use; p-value < 0.001, coefficient test for 'Random-Target' in Regression [4]). In Appendix C.1, we also show that randomized targets induced workers to achieve each of the three potential targets as often as if the target value was known.

We thus reject our EUT-Hypothesis-*i* according to which the Random-Target treatment cannot outperform the best-performing known target. The positive effect of the Random-Target treatment is robust to comparing distributions across treatments (see Section A.1.1 in Appendix A) and to conducting regressions without individual controls (see Table A.1).

Table 1. Production, internet use and randomized incentives

Dependent Variable:	*Production (std)*		*Internet use (std)*	
	[1]	[2]	[3]	[4]
Intercept	-0.256***	-0.059	0.113	-0.110
	(0.084)	(0.098)	(0.100)	(0.119)
Random-Target	0.736****	0.522****	-0.761****	-0.521****
	(0.116)	(0.131)	(0.108)	(0.132)
Known-Target Low (13)		-0.457****		0.144
		(0.134)		(0.169)
Known-Target High (39)		-0.245		0.503**
		(0.154)		(0.196)
Individual controls	Yes	Yes	Yes	Yes
p-vales (coefficient tests)				
Random-Target = Known-Target Low		<0.001		<0.001
Random-Target = Known-Target High		<0.001		<0.001
N	248	248	248	248
F	13.496****	14.178****	5.762****	5.431****
R^2	0.331	0.355	0.165	0.197
Average VIF predictors	1.224	1.292	1.224	1.292

This table reports the results from linear regressions with robust standard errors (in parentheses). (std) refers to standardized variables. Dependent variable: *Production (std)* in regressions [1] and [2], and *Internet use (std)* in regressions [3] and [4]. Individual controls include ability, cognitive skills, risk attitudes, personality and demographics (see Table A.1 in Appendix A for the complete version). Random-Target (Known-Target Low) [Known-Target High] is a dummy taking value one for the Random-Target treatment (Known-Target treatment when the target value is 13) [Known-Target treatment when the target value is 39]. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001.

5.1.2. Field data

In Figure 2, we observe that the Random-Target treatment led to higher average production levels (59.90 sums) than in all the Known-Target treatments (51.53, +16.3%, *p*-value = 0.039, coefficient test for 'Random-Target' in Regression [1] in Table 2). Similarly, the Random-Piece Rate treatment led to higher production levels (61.84) than in all the Known-Piece Rate treatments (54.95, +12.6%, *p*-value = 0.001, coefficient test for 'Random-Piece Rate' in Regression [3] in Table 2).

Furthermore, we observe that the Random-Target treatment led to a higher level of production than any of the known targets. It outperformed the low and high targets (51.35 and 48.35, *p*-values = 0.075 and 0.007, coefficient tests for 'Random-Target = Known-Target Low' and 'Random-Target = Known-Target High' in Regression [2]). It also increased production with respect to the best-performing known target (55.20) although not significantly so (*p*-value = 0.412, coefficient test for

'Random-Target' in Regression [2]). As was the case for the lab experiment, we show that randomized targets induced workers to achieve each of the three potential targets as often as if the target value was known (see Appendix C.1).

Similar results are obtained for piece rate schemes as the Random-Piece Rate treatment induced a higher level of production than the low and high piece rates (47.66 and 57.66, p-values <0.001 and 0.013, coefficient tests for 'Random-Piece Rate = Known-Piece Rate Low' and 'Random-Piece Rate = Known-Piece Rate High' in Regression [4]) as well as the medium piece rate (59.20, p-value = 0.003, coefficient test for 'Random-Piece Rate' in Regression [4]). The positive effect of the Random treatments is robust to comparing distributions (see Section A.1.2 in Appendix A) and to conducting regressions without individual controls (see bottom panel of Table A.2).

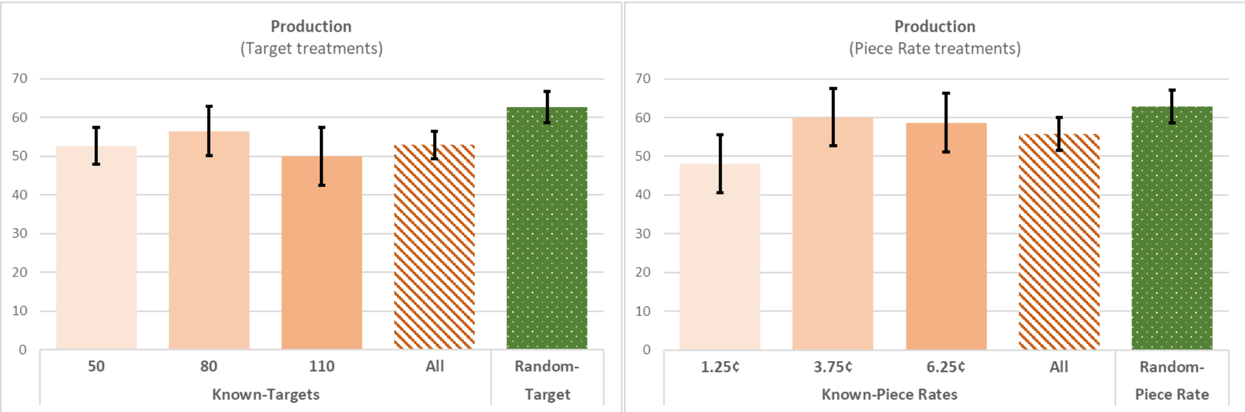


Figure 2. Average production levels for Target (left panel) and Piece Rate (right panel) treatments, with 95% confidence intervals.

The longitudinal nature of our design generated some attrition as 76.4%, 70.4%, and 62.9% of the people who attended Week 1 session completed the second, third and fourth sessions, respectively. Attrition rates were similar across treatments.[16] That said, we control for the potential effect of attrition using Lee bound estimates (Lee, 2009), which are provided in the last row of Table 2. Results are consistent with previous estimates as the effect of Random treatments is always positive except for regression [4] where the lower bound of the interval is negative and significant.

Finally, we find that the positive effect of randomized incentives held over time (see Figure 3). Indeed, average production was consistently higher across weeks under random targets [piece rates] (60.42, 63.01, 61.27 and 66.61 tables) [55.91, 62.93, 68.88 and 67.53] than under known targets [piece rates] (54.33, 53.77, 50.30 and 52.70) [51.87, 56.72, 59.20 and 56.43], and the significance of these

[16] We calculate the average attrition rate as the average proportion of workers who did not come back in a given week having participated in Week 1. Using proportion tests, we report no significant differences between Target and Piece Rate treatments (29.6% vs 30.6%, p-value = 0.806), Known- and Random-Targets (31.4% vs 28.4%, p-value = 0.586), and Known- and Random-Piece Rates (27.1% vs 32.9%, p-value = 0.289).

differences seemed to increase over time (p-values = 0.208, 0.089, 0.044 and 0.035, Mann-Whitney-Wilcoxon tests) [p-values = 0.182, 0.109, 0.097 and 0.051]. The regression analysis in Table A.3 (Appendix A) shows that the interaction coefficients for 'Random-Target × Week Number', 'Random-Piece Rate × Week Number', and 'Random Treatment × Week Number' in regressions [1], [2], and [3] are positive though not significant. This shows that the effect of randomized incentives does not decrease over time. This is an important result given the recent findings of Imas et al., (2021) who showed that lotteries triggering regret motives only increased ticket purchases the first time they were used. In our case, regret motives have a sustained effect over the four weeks of the experiment. One potential explanation is that we do not give workers an explicit regret-free alternative until the final week. As is shown in Section 6.2, people will, in line with Imas et al., (2021), refuse the regret-prone randomized schemes and opt for known schemes when offered the choice.

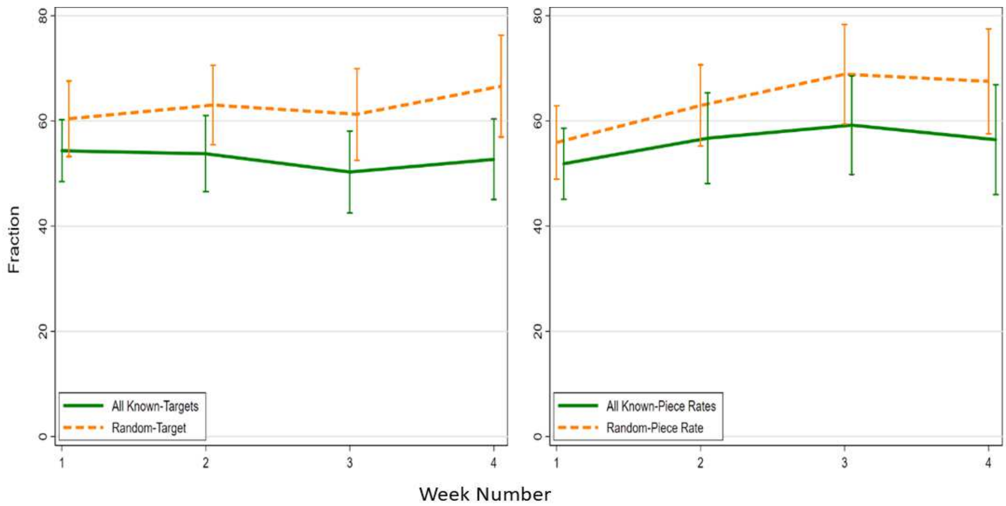


Figure 3. Average production levels by session for Target (left panel) and Piece Rate (right panel) treatments, with 95% confidence intervals.

We also analyze the impact of Random treatments pooling the data from Target and Piece Rate treatments in regression [5] in Table 2. The dummy varable 'Random Treatment', which takes value one for the Random-Target or the Random-Piece Rate treatments is positive and significant (p-value = 0.008). This shows that Random treatments significantly outperform the most productive known incentive schemes (Known-Target (80) and Known-Piece Rate (3.75)). Overall, both lab and field data concur to reject EUT-Hypothesis-*i*.

Table 2. Production and randomized targets and piece rates

Treatment:	Target Treatments		Piece Rate Treatments		Target &
	[1]	[2]	[3]	[4]	Piece Rate [5]
Intercept	-0.258**	-0.126	-0.294*	-0.270	-0.224**
	(0.125)	(0.141)	(0.161)	(0.172)	(0.113)
Random Treatments	0.224**	0.099	0.351****	0.332***	0.221***
(Random-Targets or Random-Piece Rates)	(0.109)	(0.120)	(0.102)	(0.113)	(0.084)
Known-Target Low (50)		-0.101			-0.073
		(0.090)			(0.081)
Known-Target High (110)		-0.265***			-0.235**
		(0.101)			(0.097)
Known-Piece Rate Low (1.25)				-0.116	-0.139
				(0.100)	(0.092)
Known-Piece Rate High (6.25)				0.064	0.055
				(0.073)	(0.069)
Individual controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
p-vales (coefficient tests)					
Random-Target = Known-Target Low		0.075			0.002
Random-Target = Known-Target High		0.007			0.000
Random-Piece Rate = Known-Piece Rate Low				0.000	0.001
Random-Piece Rate = Known-Piece Rate High				0.013	0.077
N	654	654	656	656	1310
χ^2	81.074****	98.083****	166.975****	169.583****	150.063****
R^2	0.174	0.177	0.327	0.331	0.204
Average VIF predictors	1.762	1.821	1.776	1.809	1.713
LEE BOUNDS INTERVALS					
(Random Treatments)					
Lower bound	0.292***	-0.158	0.178*	-0.297**	-0.227**
	(0.091)	(0.128)	(0.097)	(0.141)	(0.093)
Upper bound	0.300****	0.491****	0.276**	0.553****	0.512****
	(0.085)	(0.126)	(0.107)	(0.119)	(0.084)
Trimming proportion	0.004	0.269	0.037	0.281	0.272

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. *Individual controls* include ability, cognitive skills, risk attitudes, personality and demographics (see Table A.2 in the Appendix for the complete version). Known-Target Low (Known-Target High) [Known-Piece Rate Low {Known-Piece Rate Low} is a dummy taking value one for the Known-Target treatment when the target value is 50 (Known-Target treatment when the target value is 110) [Known-Piece Rate treatment when the target value is 1.25] {Known-Piece Rate treatment when the target value is 6.25}. Random-Target <Random-Piece Rate> is a dummy taking value one for the Random-Target <Piece Rate>Treatment. Random Treatments is a dummy variable that takes value one for the Random-Target (Random-Piece Rate) {Random-Target and Random-Piece Rate} treatments in regressions (1) and (2) [(3) and (4)] {(5)}. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001.

5.2. Choice between randomized and known schemes

In the field experiment, we conducted a final session (Week 5) in which two thirds of the workers had the possibility to choose between knowing the incentive scheme (target or piece rate) ex ante, as in the Known treatments or ex post, as in the Random treatments. In total, 122 [112] workers chose between knowing the target (50, 80 or 110) [piece rate (1.25, 3.75 or 6.25)] before or after completing the task. The vast majority of workers, 83.2%, chose to know the incentive before completing the task. The proportion who chose to know the target ex ante (85.5%) was not significantly different than the proportion who chose to know the piece rate ex ante (80.8%; *p*-value = 0.360, proportion test). Workers overwhelmingly opted for the Known treatment whether they had been assigned to bonuses (84.5%) or piece rates (82.0%, *p*-value = 0.627, proportion test), and whether they had previously competed a Known (88.0%) or a Random treatment (80.6%, *p*-value = 0.166, proportion test).

Table 3. Choice of known incentive scheme and risk attitudes

Known Scheme Dummy	[1]	[2]	[3]
Intercept	1.657****	1.310****	1.731****
	(0.436)	(0.396)	(0.418)
Target Treatment	-0.049	0.160	-0.106
	(0.243)	(0.217)	(0.236)
Risk aversion index (std)	-0.224*		-0.222*
[number of safe choices, Holt & Laury, 2002]	(0.129)		(0.132)
Willingness to take risk index (std)		-0.183	-0.213
		(0.169)	(0.185)
Individual controls	*Yes*	*Yes*	*Yes*
N	186	214	186
χ^2	38.163***	27.774	37.434**
Pseudo R^2	0.161	0.119	0.172
Average VIF predictors	1.775	1.831	1.832

This table reports the results from probit regressions with robust standard errors (in parentheses). Dependent variable: *Known Scheme Dummy*, which takes value one when a worker opted for the known scheme. *Individual controls* include ability, cognitive skills, personality and demographics (see Table A.4 in the Appendix for the complete version). Target Treatment is a dummy that takes value one for Target treatments. (std) refers to standardized variables. **p*<0.10, ***p*<0.05, ****p*<0.01 and *****p*<0.001

In Table 3, we use as dependent variable a dummy that takes value one if a person has chosen to know the exact incentive scheme (target or piece rate) ex ante rather than ex post ('Known Scheme Dummy'). Under EUT, we expect that more risk-averse workers will be more reluctant to opt for a randomized scheme. We assessed risk attitudes calculating a 'Risk aversion index' that counts the number of safe choices in the Holt and Laury (2002) elicitation task. This task was performed in Week 2 so that 78.2% of all the workers attending Week 5 took it. We also used a general risk measure (Dohmen et al., 2011) that asks people their overall willingness to take risks in a scale from 0 to 10, which was elicited in Week 0 and thus available for all workers.[17]

EUT-Hypothesis-*ii* predicts that the sign associated with the 'Risk aversion index' ('Willingness to take risk index') should be positive (negative) and significant. However, the coefficient for 'Risk aversion index' is negative and significant, and the coefficient for 'Willingness to take risk index' is negative but not significant in regressions [2] (p-value = 0.278) and [3] (p-value = 0.251). These results are inconsistent with EUT-Hypothesis-*ii* and highlight that basic measures of risk-aversion cannot explain workers' decisions to opt for the known scheme.

Next, we study behavioral mechanisms that can help explain the effectiveness of randomized schemes and the choice of incentive scheme in Week 5.

5.3. Behavioral mechanisms

5.3.1. Regret

To test the Regret-Hypothesis, we evaluate the explanatory power of regret motives on the performance of randomized schemes. A critical challenge is to quantify regret motives. From a decision theoretic viewpoint, the difficulty is to separately estimate a regret functional for a theory that allows for violations of transitivity (see Bleichrodt, Cillo and Diecidue, 2010). Because our experiment was conducted online and involved a non-student population, the elicitation task in Bleichrodt, Cillo and Diecidue (2010) proved to be excessively long and complex.[18] An appealing alternative strategy is to elicit regret motives using a psychometrically validated scale (see Marcatto and Ferrante, 2008). Using this approach, we developed a procedure to assess both monetary and effort regret in Week 1. This elicitation procedure describes a fictitious scenario that is based on the Random-Target treatment in the field experiment in which workers are placed in the role of workers (see Supplementary Material, Field Study-Additional Instructions, Week 1). We then elicit the reaction of workers to three different scenarios using a 7-item scale for five questions similar to those used in Marcatto and Ferrante (2008) such as "I wish I had made a different choice" or "I

[17] As expected, these two measures correlate negatively (ρ = -0.219, p-value < 0.001).

[18] The authors conducted individual interviews of 55 minutes on average to collect the data at an elite university in Spain. Inconsistency rates in the original studies, measured using repeated choices, ranged from 13.4% and 28.3% depending on the measurements. Our online workers recruited from an adult population in the US completed the task (including the instruction phase) in an average (median) time of only 7 (5) minutes. Perhaps not surprisingly most workers (92.1%) exhibited inconsistent choices.

am sorry about what happened to me". In the monetary regret scenario, workers were told: "Imagine that you are done with the 10-minute task, and you have completed **79** sums correctly and the target turned out to be **80**." In this case, workers will exhibit monetary regret because they have missed out on a bonus that was achievable. In the effort regret scenario, workers were told: "Imagine that you are done with the 10-minute task, and you have completed **79** sums correctly and the target turned out to be **50**." In this case, workers will exhibit effort regret because they could have worked less to receive the bonus had they known the target.[19] The reliability of the scale measured in each of the scenarios –monetary regret and effort regret– was acceptable (Cronbach α's = 0.704 and 0.837). We find that monetary regret is more pronounced than effort regret as 83.4% of the workers reported a more than neutral (a score above 4 in the 7-item scale) level of monetary regret compared to 59.4% for effort regret (p-value < 0.001, sign rank test). Monetary and effort correlate positively and significantly ($\rho = 0.247$, p-value < 0.001), which is not surprising given that they both capture related dimensions of a regretful personality (Schwartz et al., 2002).

Table 4. Production, regret and randomized incentives

Production (std)	Random Treatments [1]	Known Treatments [2]	Random & Known Treatments [3]
Intercept	-0.086	-0.156	-0.266**
	(0.171)	(0.132)	(0.113)
Regret motives			
Monetary regret (std)	0.165***	-0.008	0.001
	(0.058)	(0.069)	(0.063)
Monetary regret (std) × Random Treatments			0.150*
			(0.081)
Effort regret (std)	-0.096	-0.023	-0.068
	(0.073)	(0.093)	(0.075)
Effort regret (std) × Random Treatments			-0.021
			(0.085)
Treatment dummies			
Random Treatments			0.245***
			(0.078)
Piece Rate Treatment	0.043	-0.031	0.033
	(0.121)	(0.099)	(0.079)
Individual controls	*Yes*	*Yes*	*Yes*
N	517	701	1218

[19] For the sake of completeness, we also considered the third possible case in which people were told they the target "turned out to be **110**." In this situation, workers could exhibit both monetary (see [1_1]) and effort regret (see [1_3]) depending on whether they perceived the target as achievable or not. We do not consider this noisier measure in our analyses (Cronbach α's = 0.710).

Chi2	85.294****	76.802****	134.614****
R^2	0.218	0.239	0.214
Average VIF predictors	1.759	1.892	1.835

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. *Individual controls* include ability, cognitive skills, personality and demographics (see Table A.5 in the Appendix for the complete version). Random (Known) Treatments is a dummy variable that takes value one for the Random (Known) Target and Random (Known) Piece Rate treatments. Piece Rate Treatments is a dummy variable that takes value one for the Piece Rate treatments. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001

In Table 4, we show that monetary regret increases production levels in Random treatments (p-value = 0.005, coefficient test for 'Monetary regret' in regression [1]) but not in Known treatments (p-value = 0.905, coefficient test for 'Monetary regret' in regression [2]). Effort regret decreases production levels in Random treatments, although not significantly so (p-value = 0.188, coefficient test for 'Effort regret' in regression [1]). The size of the standardized coefficient for 'Monetary regret' is about twice as large as the coefficient for 'Effort regret', thus showing that the impact of monetary regret is substantially more pronounced. Furthermore, monetary regret amplifies the positive impact of the Random treatment on production levels (p-value = 0.063, coefficient test for 'Monetary regret × Random Treatments' in regression [3]). However, we do not find significant evidence that effort regret explains the impact of the Random treatment (p-value = 0.808, coefficient test for 'Effort regret × Random Treatments'). These findings are in line with Regret-Hypothesis *i* and *ii*.

So far, we have used measures that are psychometrically validated yet non-incentivized. In Week 5, we were able to elicit an incentivized measure of effort regret which was randomly assigned to approximately one third of the workers (126 out of 360). Incentivizing effort regret is particularly relevant as it is knowingly difficult to precisely evaluate the mental cost of effort in a given task without actually undertaking it and triggering specific physiological reactions (Boksem and Tops, 2008). In this task, we elicited the certainty equivalent for a random bonus giving $3 with probability 70% if workers completed a given target number of summations, and nothing otherwise. We measured effort regret by comparing workers' certainty equivalents when the target was equal to 1 and when it was either 40 or 80. When the target is equal to 1, all workers can achieve it without effort so that effort regret cannot play a role. In that case, the certainty equivalent for the random bonus only captures risk attitudes. By contrast, workers who would miss the target in the random bonus scheme when the target was 40 might expend a substantial level of effort in vain, consequently experiencing effort regret. Indeed, when the target was 40, the vast majority of workers could expect achieving it so that not receiving the bonus would most likely occur because the random bonus was not paid in 30.0% of the cases, thus triggering effort regret.[20] In the case in which the target is 80, the measure of effort regret is less precise because failing to receive the bonus might also be due to insufficient ability.[21] In that case, some workers might decide to exert no effort and thus exhibit no effort regret.

To measure effort regret, we define the *effort-regret premium 40* (*80*) as the difference between the random-bonus certainty equivalent when the target is 40 (80), and when it is 1.[22] The effort-regret premium captures the mental cost of exerting a substantial amount of effort while risking not being paid for it. Because effort-regret premium 40 and 80 are substantially correlated (ρ = 0.498, p-value < 0.001), we constructed a single effort-regret premium measure as the average of the two standardized measures.

In line with Regret-Hypothesis *i*, the effort-regret premium negatively impacted production levels in the Random treatment (p-value < 0.001, coefficient test for 'Effort-regret premium' in regression [1] in Table A.6 in Appendix A) but not in the Known treatment (p-value = 0.224, regression [2]). The difference in the effect of the effort-regret premium across treatments is significant (p-value = 0.009, coefficient test for 'Effort-regret premium × Random Treatments' in regression [3] in Table A.6).

Finally, in line with Regret-Hypothesis *iii*, we also show that monetary regret leads people to opt more often for the Known treatment (p-value = 0.017, coefficient test for 'Monetary regret' in regression [1] in Table A.7) whereas effort regret does not impact the choice of incentive schemes (p-value = 0.658).[23] The size of the standardized coefficient for 'Monetary regret' is about four times as large as the coefficient for 'Effort regret', again showing that monetary regret is more impactful than effort regret in explaining workers' behavior.

The fact that workers largely opted for known schemes (83.2%) in Week 5 despite producing more under randomized schemes in the previous four weeks is consistent with the existence of mental

[20] Most workers were able to achieve a target of 40 given that across all weeks and treatments 72.5% had been able to do so.
[21] A minority of workers will be able to achieve a target of 80 given that across all weeks and treatments only 31.5% had been able to do so.
[22] To calculate the random-bonus certainty equivalent, we use the midpoint of the first value of the deterministic bonus that is accepted by the worker and the last rejected value.
[23] Note that we cannot use the incentivized measure of effort regret in this regression because we elicited this measure for workers who did not make the decision across incentive schemes.

costs associated with randomized schemes. These costs become salient when directly comparing the latter with known schemes. We can provide a rough estimate of these mental costs. For example, the difference in earnings between the Random- and Known-Piece Rate was on average \$0.46 per 10 minutes of work. It follows that 80.8% of the workers were willing to forego an estimated gain of \$2.75 per hour to avoid the mental costs associated with the Random-Piece Rate.[24] This amount is similar to the reported pay of MTurkers (Hara et al., 2018), and about 40% of US minimum wage, which is the recommended pay level for online workers (Buhrmester et al., 2018; Aguinis et al., 2021). Given that risk and loss attitudes did not explain production in the Random-Piece Rate treatment and the decision to opt for the Known treatment in Week 5, we conjecture that a large share of the estimated mental costs are related to regret. An alternative estimate of these costs can be done focusing only on workers who were assigned to a Piece Rate treatment in Weeks 1 to 4 and had to choose between Random- and Known-Piece Rates. In that case, 82.9% of the workers opted for the known scheme despite an estimated additional gain associated with the Known-Piece Rate over the Random-Piece Rate of \$5.16 per hour.

5.3.2. Alternative explanations

Probability distortions

Recent research has shown that probability weighting can play a major role in incentive setting (González-Jiménez, 2021; Corgnet et al., 2023). In our Random-Target treatment, optimal production levels will critically depend on the probability weights assigned to $\frac{1}{3}$ and $\frac{2}{3}$. Empirical estimates have shown evidence of likelihood insensitivity such that $\frac{2}{3}$ is substantially underweighted ($w\left(\frac{2}{3}\right) < \frac{2}{3}$) whereas $\frac{1}{3}$ is not (see e.g., Tversky and Kahneman, 1992; Wu and Gonzalez, 1996; Prelec, 1998; Bruhin et al., 2008). That is, $w(\frac{1}{3}) \approx \frac{1}{3}$ and $w\left(\frac{2}{3}\right) < \frac{2}{3}$. Probability distortions could thus increase production levels due to the *certainty effect* where workers decide to exert a high level of effort to make sure they receive the bonus with certainty instead of with probability $\frac{2}{3}$. The positive impact of probability distortions thus relies on the magnitude of the difference: $1 - w\left(\frac{2}{3}\right)$. Under probability distortions, it follows that H-workers are more likely to

[24] For target treatments, workers' earnings did not significantly differ between the Known- and Random-Target treatments despite substantial differences in production levels. This is the case because bonus schemes are non-linear, thus making the relationship between production and earnings levels more intricate.

produce τ_H due to the certainty effect triggered by the randomized scheme. Regarding L-workers, randomization will trigger a *possibility effect*, which could lead them to produce τ_L rather than nothing, thus securing a $\frac{1}{3}$ chance of receiving the bonus. However, the possibility effect, which is captured by: $w\left(\frac{1}{3}\right) - \frac{1}{3}$, is likely to be small. Finally, M-workers can also be affected by probability distortions under the randomized scheme. These workers evaluate the difference in likelihood between $\frac{2}{3}$ and $\frac{1}{3}$ in order to decide whether to complete target τ_M. For this to happen, workers would need to exhibit a probability weighting function such that $w\left(\frac{2}{3}\right) - w\left(\frac{1}{3}\right) > \frac{1}{3}$. However, the literature has reported extensive evidence of likelihood insensitivity: $w\left(\frac{2}{3}\right) - w\left(\frac{1}{3}\right) < \frac{1}{3}$. In sum, probability distortions might induce H-workers to produce up to τ_H whereas the impact on L-workers might be limited, and they could even have a negative impact on M-workers.

Under piece rates, the incentive compatibility constraint can be written as: $e_{i,w}^{R*} = \left(\frac{w(\frac{1}{3})(p_L+p_M+p_H)}{k}\right)^{\frac{1}{\gamma_i}}$. Given that there is limited evidence of probability weighting associated with probability $\frac{1}{3}$, that is $w(\frac{1}{3}) \approx \frac{1}{3}$, we do not expect a substantial effect of probability distortions under randomized piece rates.

To test whether probability weights can explain our findings, we elicited the probability weights for our two relevant probabilities (33% and 67%) using the non-parametric method of Kpegli et al., (2023). The elicitation procedure was conducted for all the workers in the final week.[25] In line with the literature (see e.g., Tversky and Kahneman, 1992; Wu and Gonzalez, 1996; Prelec, 1998; Bruhin et al., 2008), we found limited (although significant) probability distortions for p = 33% (p-value < 0.001, mean individual estimate for $w(0.33) = 0.298$, sign rank test) and substantial underweighting for p = 67% (mean individual estimate for $w(0.67) = 0.371$, p-value < 0.001, sign rank test), which is line with the results in Camerer and Ho (1994) and Gonzalez and Wu (1999).

[25] Out of 360 workers who participated in the final week, 25 failed the attention checks and are excluded from the analyses in line with our pre-registration. The 335 remaining workers differed from our total pool of participants of 648 in terms of some individual characteristics. They were more able on the task (p-value = 0.002, rank sum test) and had higher cognitive skills (p-values < 0.001 for Raven and CRT, rank sum tests). They also reported being less likely to take risks (p-value = 0.001), being more open to experience, conscientious and more agreeable (p-values < 0.001, 0.002 and 0.017). Note that we could not reliably estimate probability weights for 5% of the workers in the final session (n = 18) who reported the same certainty equivalent for all ten decisions.

In Table A.8, we check for the presence of a possibility effect by testing whether the coefficient for '$w(0.33)$' is positive and significant in explaining production, and for the presence of a certainty effect by testing whether '$w(0.67)$' is negative and significant. None of the coefficients for probability weights are significant whether we consider Random-Target (see regressions [1] and [2]) or Random-Piece Rate treatments (see regressions [3] and [4]).[26] We also tested whether the certainty (possibility) effect was significant for high-ability (low-ability) workers conducting additional regressions including '$w(0.67)$ × Task Ability' and '$w(0.33)$ × Task Ability' as regressors (see bottom panel 'INTERACTION' in Table A.8). However, these interaction terms are not significant. Finally, in Table A.7 (regressions [2] and [3]), we also found that probability weights do not significantly explain workers' decisions to opt for the known incentive scheme.

Loss aversion and disappointment

To introduce loss aversion in our analysis, we stipulate a fixed reference point. Under bonuses, this reference value could be set at the level of the monetary bonus B so that not achieving the target will trigger losses as in DellaVigna and Pope (2018). In our experiments, we can interpret this bonus reference point as the status quo, which has been found to be empirically prevalent (see Baillon et al., 2020). This reference point is likely to be especially salient because we calibrated bonuses so that workers' compensation was aligned with standard payments for a similar task on the online platform. Under a randomized target scheme, the worker thus maximizes the following value function:

$$E[v_{i,\lambda}(e_i)] \coloneqq \frac{1}{3}\sum\nolimits_{\tau\in\Gamma}(1_{e_i\geq\tau} - \lambda 1_{e_i<\tau})B - C_i(e_i)$$

It follows that loss aversion increases the cost of not achieving the target thus inducing workers to increase their production levels. Loss aversion can lead workers to exert higher effort levels and generate higher revenues for the employer. This will hold as long as people do not have outside options that would be preferred to loss-inducing contracts. However, even when loss-free outside options are available, workers may still choose compensation contracts that result in losses because they do not fully anticipate the negative consequences of such losses (see Imas et al., 2017). Although loss aversion can impact performance, we must emphasize that, in our setting, the positive impact of a bonus reference point could be observed for both known and randomized bonus

[26] We do not introduce '$w(0.33)$' and '$w(0.67)$' as independent variables in the same regression to avoid multicollinearity issues ($\rho = 0.654$, p-value < 0.001).

schemes. Indeed, if we posit a sufficiently high level of loss aversion then setting a target τ_H might lead workers to produce at that level, regardless of their ability.

Under a randomized piece rate scheme, we can consider that workers maximize a piecewise linear value function (see e.g., Fehr and Goette, 2007) as follows: $v_{i,\lambda}(e_i) \coloneqq \frac{1}{3}\big((p_M e_i - f) + (p_H e_i - f) - \lambda(f - p_L e_i)\big) - C_i(e_i)$, where f is an exogenous reference point so that losses are only incurred when the low piece rate applies. It follows that the optimal level of production $e_{i,\lambda}^{f*} = \left(\frac{p_M + p_H + \lambda p_L}{3k}\right)^{\frac{1}{\gamma}}$ increases in loss aversion. However, the same positive effect of loss aversion can be obtained for a known piece rate scheme if we consider a reference point that is typically higher than workers' revenues ($p_i e_i$).

Although regret induces a stochastic reference point and leads to different predictions for known and randomized incentive schemes, loss aversion with a fixed reference point enhances work performance regardless of the underlying scheme (see Luft, 1994; Hannan et al., 2005; Pokorny, 2008; Berger and Pope, 2011; Sloof and van Praag, 2010; Armantier and Boly, 2015; Imas et al., 2017; Corgnet and Hernán-González, 2019; Fryer et al., 2022).

Loss aversion can potentially explain an increase in production levels in both Known and Random treatments, which limits its appeal as a distinct explanation for the positive impact of randomized incentives on performance. Furthermore, we show that loss aversion does not explain workers' performance in the Random treatments (see the negative and non-significant coefficient for 'Loss aversion index' in Table A.9 for both lab (see regression [3]) and field data (see regressions [1] and [2]), where loss aversion is measured following Brink and Rankin (2013). Finally, loss aversion does not explain workers' decision to opt for the known rather than the random schemes (see regression [4] in Table A.7).

Another behavioral model that relies on the existence of reference points and that could be applied to our setup is disappointment theory. Under randomized schemes, disappointment aversion (à la Gul, 1991) would imply that the weight assigned to the low piece rate (or high target) would be higher than the weight assigned to the high piece rate (or low target), thus leading to lower production than in the absence of disappointment.[27]

It is interesting to note that, in our setup, unlike the work environments with partial feedback that trigger disappointment aversion (see Gill and Prowse, 2012), workers receive complete feedback and can thus evaluate what they would have earned for alternative production levels. Our design is thus one in which regret is likely to play a more prominent role.

Curiosity

Random incentive schemes differ from known schemes because they resolve uncertainty about rewards at a later stage thus possibly triggering suspense (Ely et al., 2015) and curiosity (Loewenstein, 1994; Ruan et al., 2018). Furthermore, Shen et al., (2019) emphasize that a driving force of randomized incentives is that it triggers curiosity regarding the final resolution of uncertainty. According to this approach, workers will tend to work more under random incentive schemes that generate higher levels of uncertainty and will tend to opt for randomized rather than known schemes when asked to do so. We refer to these conjectures as the curiosity hypothesis. To test this hypothesis, we designed Curiosity treatments in which we increased the surprise associated with the resolution of uncertainty by increasing the entropy of rewards (see Friston et al., 2013, 2015, 2017a,b; Ely et al., 2015; Corgnet et al., 2020). In the Curiosity-Target (Curiosity-Piece Rate) treatment, the target was randomly picked after the production phase in a set of 61 integers (501 2-decimal equally spaced numbers) ranging from 50 to 110 (1.25 to 6.25), thus producing a level of Shannon entropy of 5.93 (8.97), which is substantially higher than the level of entropy (1.58) under the Random-Target and Random-Piece Rate treatments.

However, we found that workers did not exert more effort under the Curiosity-Piece Rate and the Curiosity-Target treatments than under the Random-Piece Rate and the Random-Target treatments (see negative and non-significant coefficient for 'Curiosity Treatment' in regressions [1] to [3] in Table A.10).

Furthermore, most workers opted for the Known treatment when they could choose between Known and Random treatments (83.2%, proportion test < 0.001). Finally, we elicited the interest and enjoyment scale using the intrinsic motivation inventory 7-item scale (Ryan, 1982) (Cronbach

[27] This implication of Gul's (1991) model follows from the disappointment aversion parameter (β) (see equation on top of p. 673) being positive, which has been shown to hold for gains and the range of probabilities considered in our experiment (see Abdellaoui and Bleichrodt, 2007).

α's = 0.905 and 0.799 for lab and field data) but found no evidence that Random treatments led to higher levels of interest on the task. In the lab data, the Random-Target and Known-Target treatments led to similar levels of interest on the task (p-value = 0.489, rank sum test). In the field data, we found no differences in interest between Random-Target and Known-Target treatments (p-value = 0.885, rank sum test), and between Random-Piece Rate and Known-Piece Rate treatments (p-value = 0.870, rank sum test). Similarly, we did not find differences between Curiosity-Target and Random-Target (p-value = 0.564, rank sum test), and between Curiosity-Piece Rate and Random-Piece Rate (p-value = 0.137, rank sum test).

6. Discussion

The incentive literature has overwhelmingly focused on deterministic schemes, possibly due to two reasons. First, scholars have shown deterministic schemes to be optimal under a broad range of conditions (see e.g., Laffont and Martimort, 2002). Second, researchers have questioned whether randomized schemes could be effectively implemented (see e.g., Arnott and Stiglitz, 1988). Furthermore, the existing literature in behavioral and health sciences has shown mixed evidence, thus supporting the idea that randomized schemes are indeed an intellectual curiosity not worth pursuing. Consequently, scholars and practitioners alike have widely recommended linking performance and rewards as clearly and predictably as possible (e.g., Vroom, 1987; Nadler and Lawler, 1977; Harter et al., 2003; Kim and Mauborgne, 2003; Rock, 2008; Latham, 2012).

In contrast to these views, our paper presents a series of lab and field experiments showing that randomized incentive schemes can be successfully implemented and persistently outperform deterministic schemes. These findings directly address longstanding concerns regarding the adoption of randomized schemes in work environments, particularly the potential resistance from workers due to the heightened risk associated with such schemes and the potential distrust in the randomization process.

By identifying regret as the key behavioral mechanism for our findings, we can make specific recommendations on how to use randomized schemes and prevent unintended effects. In particular, randomized schemes will have a positive effect on performance when monetary regret outweighs effort regret. While our findings suggest that monetary regret tends to be stronger than effort regret, compensation schemes characterized by a high probability of receiving no reward, irrespective of effort, might exacerbate effort regret and undermine the efficacy of randomized schemes.

(Dellavigna and Pope, 2018). We also show that randomized schemes can only be implemented effectively when workers lack an immediate opportunity to choose a regret-free deterministic scheme. This lack outside options often characterizes organizational settings, given that employees might not have immediate access to alternative compensation contracts. Therefore, our field experiment, which involved a repeated relationship with workers who likely had immediate access to multiple competing job postings, can be seen as a conservative estimate of the positive effect of randomized schemes on performance.

Our research highlights the importance of incorporating regret motives in the study of incentives because such motives will be relevant not only under randomized schemes but also, more broadly, whenever the relationship between effort and performance is stochastic. This includes virtually all classical models studied in the theory of incentives, ranging from individual incentives with output shocks to relative incentive schemes. Regret could also play a critical role in models studying the provision of information in dynamic contests (Ederer, 2010; Halac et al., 2017; Mihm and Schlapp, 2019). In the case of tournament incentives, our findings suggest that giving limited feedback to contestants about their competitors' ability and performance might boost revenues in the same way as randomized targets under bonus schemes.

Finally, we hope that our work will also inspire further research on the role of emotions in the incentive literature and encourage practitioners to consider the potential benefits of randomized incentive schemes.

7. References

Abdellaoui, M., & Bleichrodt, H. (2007). Eliciting Gul's theory of disappointment aversion by the tradeoff method. Journal of Economic Psychology, 28(6), 631-645.

Abraham, K. G., Haltiwanger, J. C., Sandusky, K., & Spletzer, J. R. 2018. Measuring the gig economy: Current knowledge and open issues (No. w24950). National Bureau of Economic Research.

Abraham, C., & Sheeran, P. (2003). Acting on intentions: The role of anticipated regret. British Journal of social psychology, 42(4), 495-511.

Acker, S., & Webber, M. (2016). Discipline and publish: The tenure review process in Ontario universities. In Assembling and governing the higher education institution (pp. 233-255). Palgrave Macmillan, London.

Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk research: Review and recommendations. Journal of Management, 47(4), 823-837.

Anselme, P., & Güntürkün, O. (2019). How foraging works: uncertainty magnifies food-seeking motivation. Behavioral and Brain Sciences, 42.

Armantier, O., and A. Boly (2015): "Framing of incentives and effort provision," International Economic Review, 56[3], 917-938.

Armstrong, M. (2019). Armstrong's handbook of reward management practice: Improving performance through reward. Sixth edition. Kogan Page Publishers.

Arnott, R., and J. E. Stiglitz (1988): "Randomization With Asymmetric Information," The Rand Journal of Economics, 19, 344-362.

Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective performance measures in optimal incentive contracts. The quarterly journal of economics, 109(4)[4], 1125-1156.

Baillon, A., Bleichrodt, H., & Spinu, V. (2020). Searching for the reference point. Management Science, 66(1), 93-112.

Balym, T. (2016). New Rio selection twist could leave Australia's swim in chaos. Herald Sun, March 11th.

Baron, D. P., & Myerson, R. B. (1982). Regulating a monopolist with unknown costs. Econometrica: Journal of the Econometric Society, 911-930.

Berger, J. and D. Pope (2011): "Can Losing Lead to Winning," Management Science, 57[5], 817-827.

Berger, J. and D. Pope (2011): "Can Losing Lead to Winning," Management Science, 57[5], 817-827.

Bernheim, B. D., & Whinston, M. D. (1998). Incomplete contracts and strategic ambiguity. American Economic Review, 902-932.

Besanko, D., Dranove, D., Shanley, M., & Schaefer, S. (2017). Economics of strategy. John Wiley & Sons.

Bleichrodt, H., Cillo, A., & Diecidue, E. (2010). A quantitative measurement of regret theory. Management Science, 56(1), 161-175.

Bleichrodt, H., & Wakker, P. P. (2015). Regret theory: A bold alternative to the alternatives. The Economic Journal, 125(583), 493-532.

Boakes R. A. (1977). Performance on learning to associate a stimulus with positive reinforcement. In Davis H.Hurwitz H. M. B. (Eds), Operant-Pavlovian Interactions Hillsdale, N. J.: Lawrence Erlbaum Associates.

Boksem, M. A., & Tops, M. (2008). Mental fatigue: costs and benefits. Brain Research Reviews, 59(1), 125-139.

Bolton, P., & Dewatripont, M. (2004). Contract theory. MIT press.
Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59.
Brewer, N. T., DeFrank, J. T., & Gilkey, M. B. (2016). Anticipated regret and health behavior: A meta-analysis. Health Psychology, 35(11), 1264.
Brink, A. G., & Rankin, F. W. (2013). The effects of risk preference and loss aversion on individual behavior under bonus, penalty, and combined contract frames. Behavioral Research in Accounting, 25(2), 145-170.
Brito, D. L., Hamilton, J. H., Slutsky, S. M., & Stiglitz, J. E. (1995). Randomization in optimal income tax schedules. Journal of Public Economics, 56(2)[2], 189-223.
Broughton, A., Gloster, R., Marvell, R., Green, M., Langley, J., & Martin, A. 2018. The experiences of individuals in the gig economy. HM Government.
Brownback, A., & Kuhn, M. A. (2019). Understanding outcome bias. Games and Economic Behavior, 117, 342-360.
Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. Econometrica, 78(4), 1375-1412.
Bryson, A., Freeman, R., Lucifora, C., Pellizzari, M., and V. Perotin (2012) "Paying for Performance: Incentive Pay Schemes and Employees' Financial Participation," CEP Discussion Paper, 1112.
Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. Perspectives on Psychological Science, 13(2), 149-154.
Camerer, C. F., & Ho, T. H. (1994). Violations of the betweenness axiom and nonlinearity in probability. Journal of risk and uncertainty, 8(2), 167-196.
Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. Science, 304(5674), 1167-1170.
Chang, L. L., DeVore, A. D., Granger, B. B., Eapen, Z. J., Ariely, D., & Hernandez, A. F. (2017). Leveraging behavioral economics to improve heart failure care and outcomes. Circulation, 136(8), 765-772.
Charness, G., & Levine, D. I. (2007). Intention and stochastic outcomes: An experimental study. The Economic Journal, 117(522), 1051-1072.
Chua, H. F., Gonzalez, R., Taylor, S. F., Welsh, R. C., & Liberzon, I. (2009). Decision-related loss: Regret and disappointment. Neuroimage, 47(4), 2031-2040.
Corgnet, B., Desantis, M., & Porter, D. (2018). What makes a good trader? On the role of intuition and reflection on trader performance. The Journal of Finance, 73(3), 1113-1137.
Corgnet, B., & Hernán-González, R. (2019). Revisiting the trade-off between risk and incentives: The shocking effect of random shocks?. Management Science, 65(3)[3], 1096-1114.
Corgnet, B., Hernán-González, R. Kpegli Y., & Zylbersztejn, A. (2023). Against the Odds! The Tradeoff Between Risk and Incentives is Alive and Well. Working Papers 2305, GATE, Université de Lyon.
Corgnet, B., Gaechter, S., & Hernán-González, R. (2020). Working too much for too little: stochastic rewards cause work addiction. Available at SSRN 3540225.
Davison, M. C. (1969). Preference for mixed-interval versus fixed-interval schedules. Journal of the Experimental Analysis of Behavior, 12[2], 247-252.

DellaVigna, S., & Pope, D. (2018). What motivates effort? Evidence and expert forecasts. The Review of Economic Studies, 85(2)[2], 1029-1069.

Diamond, W. D., & Loewy, B. Z. (1991). Effects of probabilistic rewards on recycling attitudes and behavior 1. Journal of Applied Social Psychology, 21(19), 1590-1607.

Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. American Economic Review, 101(2), 556-590.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the european economic association, 9[3], 522-550.

Dolan, P., & Rudisill, C. (2014). The effect of financial incentives on chlamydia testing rates: evidence from a randomized experiment. Social Science & Medicine, 105, 140-148.

Ederer, F. (2010). Feedback and motivation in dynamic tournaments. Journal of Economics & Management Strategy, 19(3), 733-769.

Ederer, F., Holden, R., & Meyer, M. (2018). Gaming and strategic opacity in incentive provision. The RAND Journal of Economics, 49(4)[4], 819-854.

Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. Journal of Political Economy, 123(1), 215-260.

Fehr, E., & Goette, L. (2007). Do workers work more if wages are high? Evidence from a randomized field experiment. American Economic Review, 97(1)[1], 298-317.

Fehr-Duda, H., & Epper, T. (2012). Probability and risk: Foundations and economic implications of probability-dependent risk preferences. Annual Review of Economics, 4(1)[1], 567-593.

Filiz-Ozbay, E., & Ozbay, E. Y. (2007). Auctions with anticipated regret: Theory and experiment. American Economic Review, 97(4), 1407-1418.

Filiz-Ozbay, E., Guryan, J., Hyndman, K., Kearney, M., & Ozbay, E. Y. (2015). Do lottery payments induce savings behavior? Evidence from the lab. Journal of Public Economics, 126, 1-24.

Forkman, B. A. (1993). Self-reinforced behaviour does not explain contra-freeloading in the Mongolian gerbil. Ethology, 94[2], 109-112.

Frederick, S. (2005). Cognitive reflection and decision making. Journal of Economic perspectives, 19(4), 25-42.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. 2017a. Active inference: a process theory. Neural Computation, 29(1):1–49.

Friston, K., Lin, M., Frith, C. D., Pezzulo, G., Hobson, J. A., & Ondobaka, S. 2017b. Active inference, curiosity and insight. Neural Computation, 29(10):2633–2683.

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. 2015. Active inference and epistemic value. Cognitive Neuroscience, 6(4):187–214.

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. 2013. The anatomy of choice: active inference and agency. Frontiers in Human Neuroscience, 7:598.

Fryer Jr, R. G., Levitt, S. D., List, J., & Sadoff, S. (2022). Enhancing the Efficacy of Teacher Incentives through Framing: A Field Experiment. American Economic Journal: Economic Policy, 14[4], 269-99.

Giorgetta, C., Grecucci, A., Bonini, N., Coricelli, G., Demarchi, G., Braun, C., & Sanfey, A. G. (2013). Waves of regret: A meg study of emotion and decision-making. Neuropsychologia, 51(1), 38-51.

Gerhart, B., Rynes, S. L. and I. S. Fulmer (2009): “Pay and Performance: Individuals, Groups, and Executives,” The Academy of Management Annals, 3(1)[1], 251-315.

Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. American Economic Review, 102(1)[1], 469-503.
Gill, D., Prowse, V., & Vlassopoulos, M. (2013). Cheating in the workplace: An experimental study of the impact of bonuses and productivity. Journal of Economic Behavior & Organization, 96, 120-134.
Gjesdal, F. (1982). Information and incentives: The agency information problem. The Review of Economic Studies, 49[3], 373-390.
Goette, L., & Stutzer, A. (2020). Blood donations and incentives: Evidence from a field experiment. Journal of Economic Behavior & Organization, 170, 52-74.
Gonzáles-Jiménez, V. (2021). Incentive contracts when agents distort probabilities. Universität Wien, Department of Economics.
Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. Cognitive psychology, 38(1), 129-166.
Good practice in tenure evaluation: Advice for tenured faculty, department chairs, and academic administrators. (2000). The American Council on Education, The American Association of University Professors, and United Educators Insurance Risk Retention Group.
Greenleaf, C., Gould, D., & Dieffenbach, K. (2001). Factors influencing Olympic performance: interviews with Atlanta and Negano US Olympians. Journal of applied sport psychology, 13(2)[2], 154-184.
Grossman, S. J., & Hart, O. D. (1983). An Analysis of the Principal-Agent Problem. Econometrica, 51[1], 7-45.
Gul, F. (1991). A theory of disappointment aversion. Econometrica: Journal of the Econometric Society, 667-686.
Haisley, E., Volpp, K. G., Pellathy, T., & Loewenstein, G. (2012). The impact of alternative incentive schemes on completion of health risk assessments. American Journal of Health Promotion, 26(3), 184-188.
Halac, M., Kartik, N., & Liu, Q. (2017). Contests for experimentation. Journal of Political Economy, 125(5), 1523-1569.
Haller, Hans. 1985. "The Principal-Agent Problem with a Satisficing Agent." Journal of Economic Behavior and Organization, 6(4)[4]: 359-79.
Halpern, S. D., Kohn, R., Dornbrand-Lo, A., Metkus, T., Asch, D. A., & Volpp, K. G. (2011). Lottery-based versus fixed incentives to increase clinicians' response to surveys. Health services research, 46[5], 1663-1674.
Hannan, R. L., V. B. Hoffman, & D. V. Moser (2005): "Bonus versus penalty: does contract frame affect employee effort?". In Experimental business research (pp. 151-169). Springer US.
Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., & Bigham, J. P. (2018, April). A data-driven analysis of workers' earnings on Amazon Mechanical Turk. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-14).
Hart, O., & Moore, J. (1999). Foundations of incomplete contracts. The Review of Economic Studies, 66(1)[1], 115-138.
Harter, J. K., Schmidt, F. L., & Keyes, C. L. (2003). Well-being in the workplace and its relationship to business outcomes: A review of the Gallup studies.
Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. Behavior research methods, 48, 400-407.
Hawkins, A. 2018. Uber will start rewarding high-performing drivers with better earnings and free college tuition. The Verge.

Herweg, F., Müller, D., & Weinschenk, P. (2010). Binary payment schemes: Moral hazard and loss aversion. American Economic Review, 100(5)[5], 2451-77.
Holmström, B. (1979). Moral hazard and observability. The Bell journal of economics, 74-91.
Holmström, B. (2017). Pay for performance and beyond. American Economic Review, 107(7)[7], 1753-77.
Hossain, T. and J. List. (2012): "The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations," Management Science, 58(12), 2151-2167.
Hurly, T. A., & Oseen, M. D. (1999). Context-dependent, risk-sensitive foraging preferences in wild rufous hummingbirds. Animal Behaviour, 58[1], 59-66.
Huws, U., Spencer, N., Syrdal, D. S., & Holts, K. 2017. Work in the European gig economy: research results from the UK, Sweden, Germany, Austria, the Netherlands, Switzerland and Italy.
Imas, A. L. E. X., Lamé, D. I. E. G. O., & Wilson, A. J. (2021). Reversals between one-shot and repeated decisions in incentive design: the case of regret. Working Paper.
Imas, A., Sadoff, S., & Samek, A. (2017). Do people anticipate loss aversion?. Management Science, 63[5], 1271-1284.
Jarzabkowski, P., Sillince, J. A., & Shaw, D. (2010). Strategic ambiguity as a rhetorical resource for enabling multiple interests. Human Relations, 63(2), 219-248.
Jehiel, P. (2015). On transparency in organizations. The Review of Economic Studies, 82(2)[2], 736-761.
John, O., Donahue, E. and R. Kentle (1991) "The Big Five Inventory—Versions 4a and 54," Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
John, O., Naumann, L. and C. Soto (2008) "Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, Measurement, And Conceptual Issues," In O. John, R. Robins, and L. Pervin (Eds.), Handbook of Personality: Theory and Research, 3, 114-158. New York, NY: Guilford Press.
Kadan, O., Reny, P. J., & Swinkels, J. M. (2017). Existence of Optimal Mechanisms in Principal-Agent Problems. Econometrica, 85(3)[3], 769-823.
Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47[2], 263-291.
Kim, W. C., & Mauborgne, R. (2003). Fair process: Managing in the knowledge economy. Harvard business review, 81(1)[1], 127-136.
Kimmel, S. E., Troxel, A. B., Loewenstein, G., Brensinger, C. M., Jaskowiak, J., Doshi, J. A., ... & Volpp, K. (2012). Randomized trial of lottery-based incentives to improve warfarin adherence. American heart journal, 164(2), 268-274.
Klein, R. A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bahník, Š., Bernstein, M. J., ... & Nosek, B. A. (2014). Investigating variation in replicability. Social psychology 45(3).
Kpegli, Y. T., Corgnet, B., & Zylbersztejn, A. (2023). All at once! A comprehensive and tractable semi-parametric method to elicit prospect theory components. Journal of Mathematical Economics, 104, 102790.
Lam, L., Cheng, B., Bamberger, P., & Wong, M. (2022). Research: The unintended consequences of pay transparency. Harvard Business Review, August 12, 2022.
Latham, G. P. (2012). Motivate employee performance through goal setting. Handbook of principles of organizational behavior: Indispensable knowledge for evidence-based management, 161-178.

Lazear, E. P. (2018). Compensation and incentives in the workplace. Journal of Economic Perspectives, 32(3)[3], 195-214.
Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. The Review of Economic Studies, 76[3], 1071-1102.
Levitt, S. D., List, J. A., & Sadoff, S. (2016). The effect of performance-based incentives on educational achievement: Evidence from a randomized experiment (No. w22107). National Bureau of Economic Research.
Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. Psychological Bulletin, 116(1), 75.
Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. Psychological bulletin, 116[1], 75.
Loomes, G. (1988). Further evidence of the impact of regret and disappointment in choice under uncertainty. Economica, 47-62.
Loomes, G., & Sugden, R. (1987). Testing for regret and disappointment in choice under uncertainty. The Economic Journal, 97, 118-129.
Luft, J. (1994): "Bonus and penalty incentives contract choice by employees," Journal of Accounting and Economics, 18[2], 181-206.
Marcatto, F., & Ferrante, D. (2008). The regret and disappointment scale: An instrument for assessing regret and disappointment in decision making. Judgment and Decision making, 3[1], 87.
Maskin, E., & Riley, J. (1984). Monopoly with incomplete information. The RAND Journal of Economics, 15(2)[2], 171-196.
McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. American Psychologist, 52(5), 509.
Mihm, J., & Schlapp, J. (2019). Sourcing innovation: On feedback in contests. Management science, 65(2), 559-576.
Milgrom, P. and Roberts, J. (1992) Economics, Organization & Management. Prentice Hall, New York.
Nadler, D. A., & Lawler, E. (1977). Motivating individuals in organizational settings. Perspectives on behavior in organizations (np). New York, New York: McGraw-Hill.
Pokorny, K. (2008). Pay—but do not pay too much: An experimental study on the impact of incentives. Journal of Economic Behavior & Organization, 66[2], 251-264.
O'Connor, E., McCormack, T., & Feeney, A. (2012). The development of regret. Journal of experimental child psychology, 111(1), 120-127.
Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. Journal of experimental social psychology, 45(4), 867-872.
Prelec, D. (1998). The probability weighting function. Econometrica, 497-527.
Prottas, D. J., Shea-Van Fossen, R. J., Cleaver, C. M., & Andreassi, J. K. (2017). Relationships among faculty perceptions of their tenure process and their commitment and engagement. Journal of Applied Research in Higher Education.
Quiggin, J. (1982), A theory of anticipated utility. Journal of Economic Behavior and Organization 3[4], 323-43.
Raven, J. C. (1936). Mental tests used in genetic studies: The performance of related individuals on tests mainly educative and mainly reproductive. MSc Thesis, University of London.
Rock, D. (2008). SCARF: A brain-based model for collaborating with and influencing others. NeuroLeadership journal, 1[1], 44-52.

Rothgeb, J. M., & Burger, B. (2009). Tenure standards in political science departments: Results from a survey of department chairs. PS: Political Science & Politics, 42(3)[3], 513-519.
Ruan, B., Hsee, C. K., & Lu, Z. Y. (2018). The teasing effect: An underappreciated benefit of creating and resolving an uncertainty. Journal of Marketing Research, 55[4], 556-570.
Rubin, J., & Sheremeta, R. (2016). Principal–agent settings with random shocks. Management Science, 62[4], 985-999.
Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. Journal of Personality and Social Psychology, 43(3), 450.
Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Varied forms of persistence: When free-choice behavior is not intrinsically motivated. Motivation and Emotion, *15*, 185-205.
Sapolsky, R. M. (2017). Behave: The biology of humans at our best and worst. Penguin.
Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: happiness is a matter of choice. Journal of personality and social psychology, 83[5], 1178.
Sen, A. P., Sewell, T. B., Riley, E. B., Stearman, B., Bellamy, S. L., Hu, M. F., ... & Volpp, K. G. (2014). Financial incentives for home-based health monitoring: a randomized controlled trial. Journal of general internal medicine, 29, 770-777.
Shen, L., Fishbach, A., & Hsee, C. K. (2015). The motivating-uncertainty effect: Uncertainty increases resource investment in the process of reward pursuit. Journal of Consumer Research, 41[5], 1301-1315.
Shen, L., Hsee, C. K., & Talloen, J. H. (2019). The fun and function of uncertainty: Uncertain incentives reinforce repetition decisions. Journal of Consumer Research, 46[1], 69-81.
Skinner, B. F. (1951). How to teach animals. Scientific American, 185[6], 26-29.
Sloof, R., & Van Praag, C. M. (2010). The effect of noise in a performance measure on work motivation: A real effort laboratory experiment. Labour Economics, 17(5)[5], 751-765.
Somasundaram, J., & Diecidue, E. (2017). Regret theory and risk attitudes. Journal of risk and uncertainty, 55(2), 147-175.
Starmer, C. (1992). Testing new theories of choice under uncertainty using the common consequence effect. The Review of Economic Studies, 59(4), 813-830.
Starmer, C., & Sugden, R. (1989). Probability and juxtaposition effects: An experimental investigation of the common ratio effect. Journal of Risk and Uncertainty, 2(2), 159-178.
Starmer, C., & Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. Journal of risk and uncertainty, 6(3), 235-254.
Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. Journal of Risk and uncertainty, 32(2)[2], 101-130.
Strausz, R. (2006). Deterministic versus stochastic mechanisms in principal–agent models. Journal of economic theory, 128(1)[1], 306-314.
Terstiege, S. (2012). Endogenous information and stochastic contracts. Games and Economic Behavior, 76(2)[2], 535-547.
Thirumurthy, H., Masters, S. H., Rao, S., Murray, K., Prasad, R., Zivin, J. G., ... & Agot, K. (2016). The effects of providing fixed compensation and lottery-based rewards on uptake of medical male circumcision in Kenya: a randomized trial. Journal of acquired immune deficiency syndromes (1999), 72(Suppl 4), S309.
Tirole, J. (1999). Incomplete contracts: Where do we stand?. Econometrica, 67(4)[4], 741-781.
Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. Thinking & reasoning, 20(2), 147-168.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and uncertainty, 5[4], 297-323.

Van Dijk, W. W., van der Pligt, J., & Zeelenberg, M. (1999). Effort invested in vain: The impact of effort on the intensity of disappointment and regret. Motivation and Emotion, 23(3), 203-220.

Volpp, K. G., Loewenstein, G., Troxel, A. B., Doshi, J., Price, M., Laskin, M., & Kimmel, S. E. (2008). A test of financial incentives to improve warfarin adherence. BMC health services research, 8(1), 1-6.

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. Management science, 42(12), 1676-1690.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. Organizational Behavior and Human Decision Processes, 93(2), 155-168.

Zeiler, M. D. (1972). Fixed-interval behavior: effects of percentage reinforcement. Journal of the Experimental Analysis of Behavior, 17[2], 177-189.

Appendices

A. Additional results

A.1. Cumulative distribution functions across treatments

A.1.1. Lab data

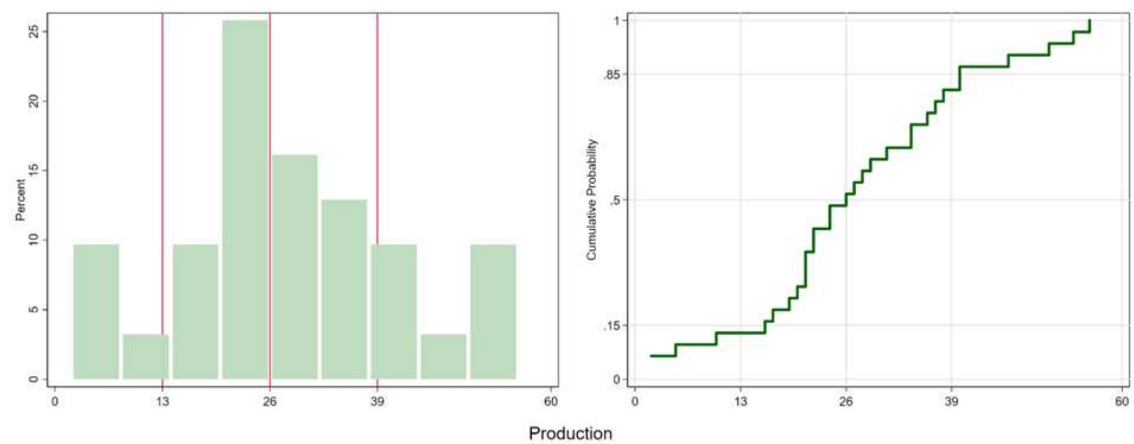
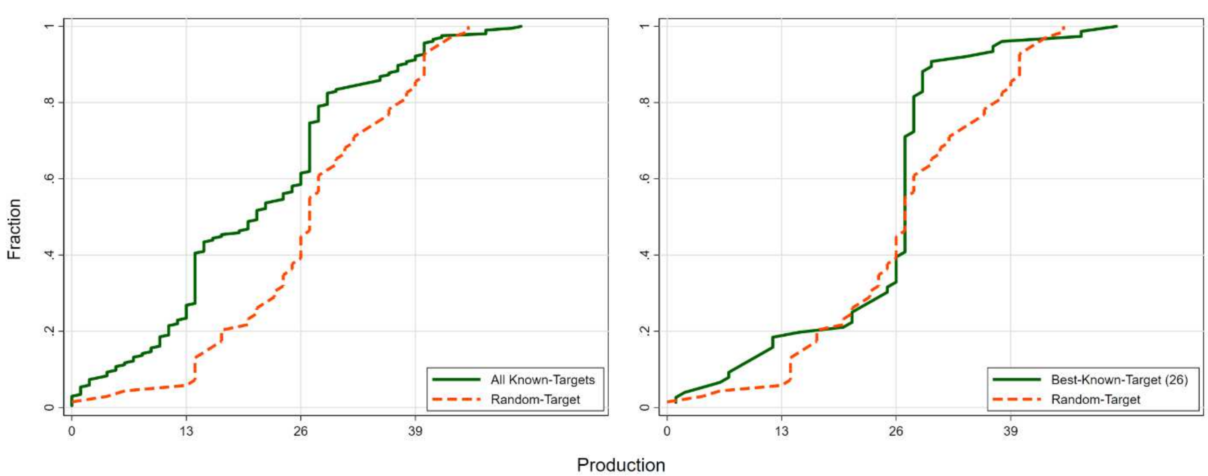


Figure A.1. Histogram (left panel) and cumulative distribution (right panel) of production levels for the calibration exercise with 31 participants paid a piece rate of 16¢ per correct answer. The median production level is 26 with a first and fourth quartile production levels of 20 and 37.

Production levels

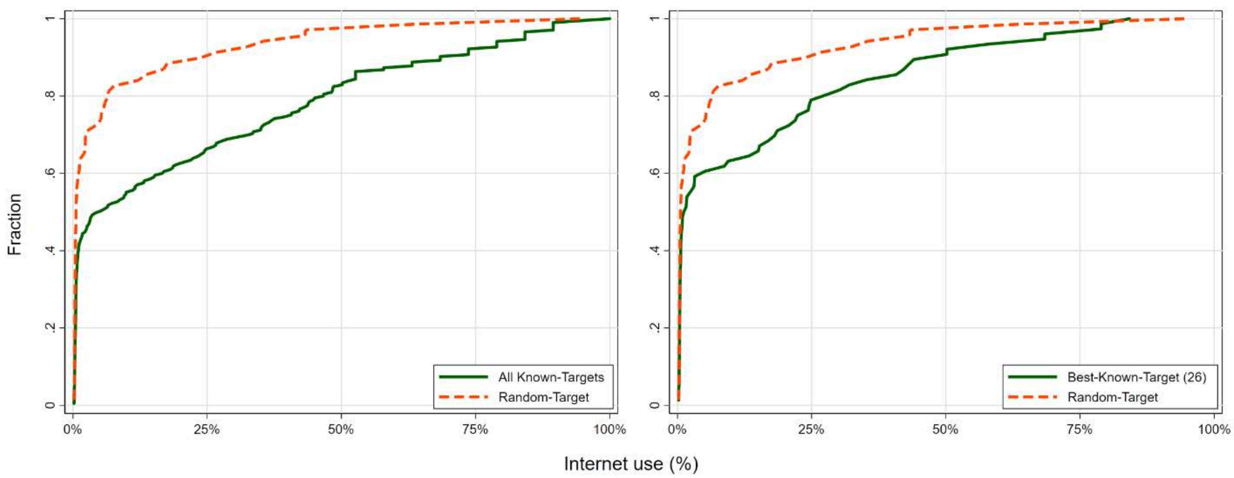


(a) Kolmgorov-Smirnov Test (p-value < 0.001) **(b)** Kolmgorov-Smirnov Test (p-value = 0.008)
p-value reported for the hypothesis: Known-Target(s) CDF above Random-Tandom CDF.[28]

Figure A.2. Cumulative Distribution Functions of production levels for **(a)** all Known-Target treatments (solid blue curve) and the Random-Target treatment (dashed red curve), and **(b)** Best-Known-Target treatment (26) (solid blue curve) and the Random-Target treatment (dashed red curve).

[28] Combined p-values lead to similar results.

Internet use (%)

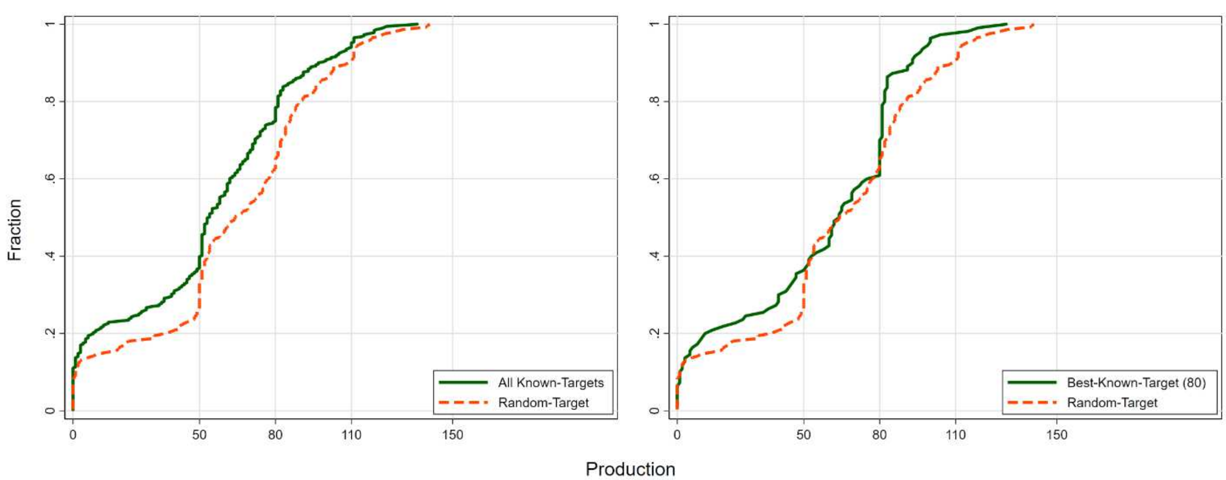


(a) Kolmgorov-Smirnov Test (p-value < 0.001) **(b)** Kolmgorov-Smirnov Test (p-value = 0.029)
p-value reported for the hypothesis: Known-Target(s) CDF above Random-Target CDF.

Figure A.3. Cumulative Distribution Functions of internet use for **(a)** all Known-Target treatments (solid blue curve) and the Random-Target treatment (dashed red curve), and **(b)** Best-Known Target treatment (26) (solid blue curve) and the Random-Target treatment (dashed red curve).

A.1.2. *Field data*

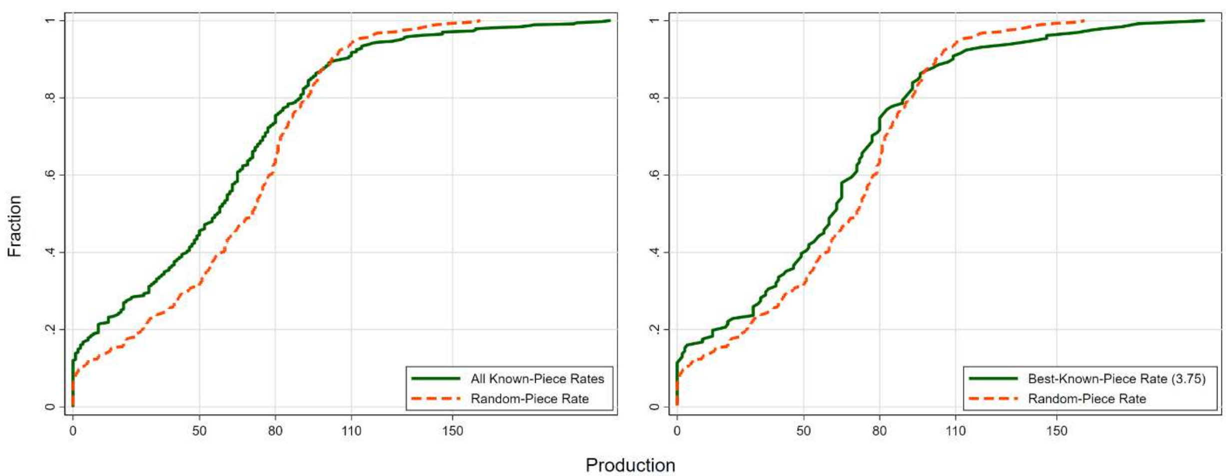
Target treatments



(a) Kolmgorov-Smirnov Test (p-value = 0.003) **(b)** Kolmgorov-Smirnov Test (p-value = 0.106)
p-value reported for the hypothesis: Known-Target(s) CDF above Random-Target CDF.

Figure A.4. Cumulative Distribution Functions of internet use for **(a)** all Known-Target treatments (solid blue curve) and the Random-Target treatment (dashed red curve), and **(b)** Best-Known Target treatment (80) (solid blue curve) and the Random-Target treatment (dashed red curve).

Piece Rate treatments



(a) Kolmgorov-Smirnov Test (p-value < 0.001)   **(b)** Kolmgorov-Smirnov Test (p-value = 0.078)
p-value reported for the hypothesis: Known-Piece Rate(s) CDF above Random-Piece Rate CDF.

Figure A.5. Cumulative Distribution Functions of production levels for **(a)** all Known-Piece Rate treatments (solid blue curve) and the Random-Piece Rate treatment (dashed red curve), and **(b)** Best-Known-Piece Rate treatment (3.75) (solid blue curve) and the Random-Piece Rate treatment (dashed red curve).

A.2. Additional Tables

Table A.1. Production, internet use and randomized incentives

Dependent Variable:	*Production (std)*		*Internet use (std)*	
	[1]	[2]	[3]	[4]
Intercept	-0.256***	-0.059	0.113	-0.110
	(0.084)	(0.098)	(0.100)	(0.119)
Treatment dummies				
Random-Target	0.736****	0.522****	-0.761****	-0.521****
	(0.116)	(0.131)	(0.108)	(0.132)
Known-Target Low (13)		-0.457****		0.144
		(0.134)		(0.169)
Known-Target High (39)		-0.245		0.503**
		(0.154)		(0.196)
Ability and cognitive skills				
Task ability (std)	0.446****	0.433****	-0.124	-0.110
	(0.067)	(0.066)	(0.080)	(0.082)
Raven (std)	0.046	0.062	-0.069	-0.079
	(0.062)	(0.060)	(0.074)	(0.072)
CRT (std)	0.070	0.068	0.056	0.061
	(0.072)	(0.072)	(0.078)	(0.081)
Risk attitudes				
Risk aversion index (std)	-0.080	-0.073	0.058	0.031
[number of safe choices, Holt & Laury, 2002]	(0.052)	(0.052)	(0.063)	(0.062)
Personality traits				
Openness (std)	-0.166***	-0.161***	0.113	0.094
	(0.060)	(0.059)	(0.076)	(0.073)
Conscientiousness (std)	0.057	0.060	-0.056	-0.064
	(0.063)	(0.061)	(0.078)	(0.075)
Extraversion (std)	-0.062	-0.069	0.096	0.083
	(0.058)	(0.056)	(0.071)	(0.068)
Agreeableness (std)	0.110*	0.117*	-0.143*	-0.132*
	(0.064)	(0.061)	(0.078)	(0.075)
Neuroticism (std)	-0.011	0.003	0.037	0.037
	(0.069)	(0.069)	(0.078)	(0.078)
Demographics				
Male Dummy	-0.136	-0.098	0.287*	0.264*
	(0.134)	(0.131)	(0.161)	(0.159)
Age (std)	0.091***	0.102***	-0.140***	-0.139***
	(0.034)	(0.032)	(0.048)	(0.045)
N	248	248	248	248
F	13.496****	14.178****	5.762****	5.431****
R^2	0.331	0.355	0.165	0.197
Average VIF predictors	1.224	1.292	1.224	1.292
NO INDIVIDUAL CONTROLS				
Random-Target	0.561****	0.257*	-0.594****	-0.310**
	(0.128)	(0.146)	(0.113)	(0.132)

This table reports the results from linear regressions with robust standard errors (in parentheses). Dependent variable: *Production (std)* in regressions [1] and [2], and *Internet use (std)* in regressions [3] and [4]. Random-Target (Known-Target Low) [Known-Target High] is a dummy taking value one for the Random-Target treatment (Known-Target treatment when the target value is 13) [Known-Target treatment when the target value is 39]. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. **p*<0.10, ***p*<0.05, ****p*<0.01 and *****p*<0.001.

Table A.2. Production and randomized targets and piece rates

Treatment: *Production (std)*	Target Treatments		Piece Rate Treatments		Target & Piece Rate
	[1]	[2]	[3]	[4]	[5]
Intercept	-0.258**	-0.126	-0.294*	-0.270	-0.224**
	(0.125)	(0.141)	(0.161)	(0.172)	(0.113)
Treatment dummies					
Random Treatments	0.224**	0.099	0.351****	0.332***	0.221***
	(0.109)	(0.120)	(0.102)	(0.113)	(0.084)
Known-Target Low (50)		-0.101			-0.073
		(0.090)			(0.081)
Known-Target High (110)		-0.265***			-0.235**
		(0.101)			(0.097)
Known-Piece Rate Low (1.25)				-0.116	-0.139
				(0.100)	(0.092)
Known-Piece Rate High (6.25)				0.064	0.055
				(0.073)	(0.069)
Ability and cognitive skills					
Task ability (std)	0.243****	0.245****	0.379****	0.378****	0.305****
	(0.058)	(0.058)	(0.080)	(0.080)	(0.056)
Raven (std)	0.153**	0.154**	0.045	0.044	0.099*
	(0.069)	(0.069)	(0.069)	(0.069)	(0.052)
CRT (std)	0.038	0.037	0.200**	0.199**	0.093*
	(0.076)	(0.076)	(0.080)	(0.080)	(0.056)
Risk attitudes					
Willingness to take risk index (std)[29]	0.213****	0.214****	0.013	0.012	0.105**
	(0.057)	(0.056)	(0.062)	(0.062)	(0.044)
Personality traits					
Openness (std)	-0.048	-0.049	-0.101	-0.098	-0.053
	(0.075)	(0.076)	(0.084)	(0.084)	(0.063)
Conscientiousness (std)	-0.031	-0.032	0.164**	0.161**	0.074
	(0.090)	(0.091)	(0.080)	(0.080)	(0.061)
Extraversion (std)	0.017	0.012	-0.135*	-0.125	-0.058
	(0.066)	(0.066)	(0.077)	(0.076)	(0.053)
Agreeableness (std)	0.197**	0.201***	-0.018	-0.016	0.079
	(0.077)	(0.077)	(0.088)	(0.087)	(0.059)
Neuroticism (std)	0.099	0.100	-0.084	-0.076	0.032
	(0.079)	(0.079)	(0.066)	(0.066)	(0.058)
Demographics					
Male Dummy	0.087	0.084	-0.234*	-0.238*	-0.036
	(0.111)	(0.112)	(0.132)	(0.132)	(0.088)
Age (std)	0.089	0.089	-0.118*	-0.114*	-0.002
	(0.057)	(0.057)	(0.066)	(0.066)	(0.045)
N	654	654	656	656	1310
χ^2	81.074****	98.083****	166.975****	169.583****	150.063****
R^2	0.174	0.177	0.327	0.331	0.204
Average VIF predictors	1.762	1.821	1.776	1.809	1.713
NO INDIVIDUAL CONTROLS					
Random Treatments	0.276**	0.155	0.257**	0.371***	0.243***
	(0.111)	(0.121)	(0.127)	(0.116)	(0.083)

[29] We use the "Willingness to take risk index" instead of the "Risk-aversion index" because it was available for all workers.

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. Known-Target Low (Known-Target High) [Known-Piece Rate Low {Known-Piece Rate Low} is a dummy taking value one for the Known-Target treatment when the target value is 50 (Known-Target treatment when the target value is 110) [Known-Piece Rate treatment when the target value is 1.25] {Known-Piece Rate treatment when the target value is 6.25}. Random-Target <Random-Piece Rate> is a dummy taking value one for the Random-Target <Piece Rate>Treatment. Random Treatments is a dummy variable that takes value one for the Random-Target (Random-Piece Rate) {Random-Target and Random-Piece Rate} treatments in regressions (1) and (2) [(3) and (4)] {(5)}. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001.

Table A.3. Production, randomized targets and piece rates, and learning

Treatments: *Production (std)*	Target Treatments [1]	Piece Rate Treatments [2]	Target & Piece Rate [3]
Intercept	-0.203	-0.300	-0.287**
	(0.149)	(0.193)	(0.129)
Treatment dummies			
Random Treatments	0.082	0.187	0.127
	(0.166)	(0.174)	(0.118)
Random Treatments × Week number	0.043	0.051	0.049
	(0.047)	(0.049)	(0.034)
Week number	-0.017	0.017	0.001
	(0.028)	(0.037)	(0.023)
Piece Rate Treatment			0.028
			(0.075)
Ability and cognitive skills			
Task ability (std)	0.241****	0.380****	0.305****
	(0.058)	(0.080)	(0.056)
Raven (std)	0.154**	0.045	0.100*
	(0.069)	(0.069)	(0.053)
CRT (std)	0.039	0.198**	0.090
	(0.076)	(0.080)	(0.056)
Risk attitudes			
Willingness to take risk index (std)	0.212****	0.011	0.103**
	(0.057)	(0.062)	(0.044)
Personality traits			
Openness (std)	-0.050	-0.100	-0.054
	(0.075)	(0.084)	(0.064)
Conscientiousness (std)	-0.030	0.163**	0.076
	(0.091)	(0.080)	(0.062)
Extraversion (std)	0.018	-0.133*	-0.059
	(0.066)	(0.077)	(0.054)
Agreeableness (std)	0.195**	-0.016	0.076
	(0.077)	(0.088)	(0.060)
Neuroticism (std)	0.099	-0.083	0.030
	(0.079)	(0.067)	(0.058)
Demographics			
Male Dummy	0.087	-0.234*	-0.032
	(0.111)	(0.132)	(0.088)
Age (std)	0.088	-0.116*	-0.002
	(0.056)	(0.066)	(0.045)
N	654	656	1310
χ^2	80.984****	177.186****	138.618****
R^2	0.174	0.327	0.200
Average VIF predictors	2.776	2.703	2.593
NO INDIVIDUAL CONTROLS			
Random Treatments × Week number	0.052	0.053	0.053
	(0.046)	(0.049)	(0.034)

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. Random Treatments is a dummy variable that takes value one for the Random-Target (Random-Piece Rate) {Random-Target and Random-Piece Rate} treatments in regressions (1) [(2)] {(3)}. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001.

Table A.4. Choice of Known incentive scheme and risk attitudes

Known Scheme Dummy	[1]	[2]	[3]
Intercept	1.657****	1.310****	1.731****
	(0.436)	(0.396)	(0.418)
Treatment dummies			
Target Treatment	-0.049	0.160	-0.106
	(0.243)	(0.217)	(0.236)
Ability and cognitive skills			
Task ability (std)	-0.218*	-0.185	-0.213
	(0.131)	(0.125)	(0.135)
Raven (std)	0.161	0.134	0.143
	(0.161)	(0.142)	(0.164)
CRT (std)	0.257*	0.188	0.242
	(0.154)	(0.134)	(0.152)
Risk attitudes			
Risk aversion index (std)	-0.224*		-0.222*
[number of safe choices, Holt & Laury, 2002]	(0.129)		(0.132)
Willingness to take risk index (std)[30]		-0.183	-0.213
		(0.169)	(0.185)
Personality traits			
Openness (std)	-0.132	-0.136	-0.176
	(0.184)	(0.170)	(0.194)
Conscientiousness (std)	0.113	0.097	0.043
	(0.196)	(0.189)	(0.214)
Extraversion (std)	0.148	0.254	0.248
	(0.157)	(0.155)	(0.178)
Agreeableness (std)	-0.094	-0.154	-0.114
	(0.160)	(0.155)	(0.162)
Neuroticism (std)	0.107	0.162	0.070
	(0.198)	(0.187)	(0.206)
Demographics			
Male Dummy	-0.071	0.036	-0.049
	(0.254)	(0.235)	(0.257)
Age (std)	-0.146	-0.131	-0.205
	(0.124)	(0.118)	(0.133)
N	186	214	186
χ^2	38.163***	27.774	37.434**
Pseudo R^2	0.161	0.119	0.172
Average VIF predictors	1.775	1.831	1.832
NO INDIVIDUAL CONTROLS			
Risk aversion index (std)	-0.150		-0.154
	(0.120)		(0.122)
Willingness to take risk index (std)		-0.073	-0.087
		(0.103)	(0.115)

This table reports the results from probit regressions with robust standard errors (in parentheses). Dependent variable: *Known Scheme Dummy*, which takes value one when a worker opted for the known scheme. Target Treatment is a dummy that takes value one for Target treatments. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. **p*<0.10, ***p*<0.05, ****p*<0.01 and *****p*<0.001

[30] We use the "Willingness to take risk index" instead of the "Risk-aversion index" in the field study because it was available for all workers.

Table A.5. Production, regret and randomized incentives

Treatments: *Production (std)*	Random Treatments [1]	Known Treatments [2]	Random & Known Treatments [3]
Intercept	-0.086	-0.156	-0.266**
	(0.171)	(0.132)	(0.113)
Regret motives			
Monetary regret (std)	0.165***	-0.008	0.001
	(0.058)	(0.069)	(0.063)
Monetary regret (std) × Random Treatments			0.150*
			(0.081)
Effort regret (std)	-0.096	-0.023	-0.068
	(0.073)	(0.093)	(0.075)
Effort regret (std) × Random Treatments			-0.021
			(0.085)
Treatment dummies			
Random Treatments			0.245***
			(0.078)
Piece Rate Treatments	0.043	-0.031	0.033
	(0.121)	(0.099)	(0.079)
Ability and cognitive skills			
Task ability (std)	0.379****	0.355****	0.340****
	(0.072)	(0.084)	(0.061)
Raven (std)	-0.066	0.201***	0.102*
	(0.085)	(0.071)	(0.056)
CRT (std)	0.190**	-0.068	0.032
	(0.089)	(0.084)	(0.063)
Risk attitudes			
Willingness to take risk index (std)	0.139**	0.058	0.098**
	(0.064)	(0.069)	(0.047)
Openness (std)	-0.029	-0.075	-0.084
	(0.102)	(0.099)	(0.070)
Personality traits			
Conscientiousness (std)	0.025	0.104	0.079
	(0.131)	(0.082)	(0.068)
Extraversion (std)	-0.026	-0.025	-0.032
	(0.078)	(0.088)	(0.057)
Agreeableness (std)	0.040	0.098	0.071
	(0.079)	(0.094)	(0.062)
Neuroticism (std)	0.068	0.061	0.038
	(0.088)	(0.081)	(0.059)
Demographics			
Male Dummy	-0.062	-0.040	-0.068
	(0.127)	(0.117)	(0.092)
Age (std)	-0.076	0.071	-0.006
	(0.060)	(0.072)	(0.047)
N	517	701	1218
χ^2	85.294****	76.802****	134.614****
R^2	0.218	0.239	0.214
Average VIF predictors	1.759	1.892	1.835

NO INDIVIDUAL CONTROLS			
Monetary regret (std)	0.181***	0.017	0.019
	(0.061)	(0.068)	(0.068)
Monetary regret (std) ×			0.153*
Random Treatments			(0.092)
Effort regret (std)	-0.200***	-0.261****	-0.252****
	(0.066)	(0.063)	(0.063)
Effort regret (std) ×			0.060
Random Treatments			(0.091)

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. Random (Known) Treatments is a dummy variable that takes value one for the Random- (Known-) Target and Random- (Known-) Piece Rate treatments. Piece Rate Treatments is a dummy variable that takes value one for the Piece Rate treatments. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001

Table A.6. Production, effort-regret premium and randomized incentives

Treatments:	Random Treatments	Known Treatments	All Treatments
Production (std)	[1]	[2]	[3]
Intercept	-0.400	-0.169	-0.423*
	(0.295)	(0.270)	(0.231)
Treatment dummies			
Effort regret premium (std)	-0.477****	0.105	0.082
	(0.087)	(0.086)	(0.094)
Effort regret premium (std) ×			-0.448***
Random Treatments			(0.171)
Treatment dummies			
Piece Rate Treatments	0.287	0.374*	0.505***
	(0.196)	(0.209)	(0.174)
Random Treatments			0.078
			(0.176)
Ability and cognitive skills			
Task ability (std)	0.759****	0.440***	0.442****
	(0.120)	(0.139)	(0.105)
Raven (std)	-0.025	0.211	0.101
	(0.143)	(0.164)	(0.116)
CRT (std)	0.448****	-0.147	0.022
	(0.133)	(0.134)	(0.110)
Risk attitudes			
Willingness to take risk index (std)	0.108	0.302**	0.130
	(0.094)	(0.153)	(0.099)
Personality traits			
Openness (std)	0.150	0.050	0.028
	(0.161)	(0.141)	(0.122)
Conscientiousness (std)	-0.824****	0.196	-0.213
	(0.203)	(0.179)	(0.155)
Extraversion (std)	0.205	0.038	-0.005
	(0.161)	(0.142)	(0.115)
Agreeableness (std)	0.457***	0.233	0.338**
	(0.166)	(0.235)	(0.144)
Neuroticism (std)	-0.212**	0.404*	0.036
	(0.083)	(0.209)	(0.154)
Demographics			
Male Dummy	0.160	-0.207	-0.019
	(0.165)	(0.255)	(0.178)
Age (std)	0.100	0.318**	0.181*
	(0.106)	(0.150)	(0.102)
N	102	213	315
χ^2	935.620****	53.338****	82.606****
R^2	0.718	0.455	0.417
Average VIF predictors	3.015	2.629	2.070
NO INDIVIDUAL CONTROLS			
Effort regret premium (std)	-0.290*	-0.130	-0.114
	(0.154)	(0.095)	(0.088)
Effort regret premium (std) ×			-0.178
Random Treatments			(0.167)

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. Piece Rate (Random) Treatments is a dummy variable that takes value one for the Piece Rate (Random) treatments. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001

Table A.7. Choice of known incentive scheme and risk attitudes

Known Scheme Dummy	[1]	[2]	[3]	[4]
Intercept	1.433****	1.427****	1.435****	1.397****
	(0.394)	(0.395)	(0.396)	(0.402)
Regret motives				
Monetary regret (std)	0.357**	0.369**	0.380**	0.375**
	(0.149)	(0.155)	(0.160)	(0.165)
Effort regret (std)	-0.090	-0.109	-0.115	-0.182
	(0.203)	(0.204)	(0.206)	(0.215)
Probability weights				
w(0.33) (std)		-0.036		
		(0.135)		
w(0.67) (std)			-0.071	
			(0.141)	
Loss aversion				
Loss aversion index (std)				0.101
				(0.134)
Treatment dummies				
Target Treatment	0.304	0.242	0.254	0.425*
	(0.233)	(0.236)	(0.238)	(0.252)
Ability and cognitive skills				
Task ability (std)	-0.289**	-0.253*	-0.247*	-0.326**
	(0.145)	(0.151)	(0.149)	(0.150)
Raven (std)	0.017	-0.021	-0.026	0.002
	(0.172)	(0.182)	(0.181)	(0.174)
CRT (std)	0.208	0.212	0.217	0.215
	(0.161)	(0.166)	(0.167)	(0.171)
Risk attitudes				
Willingness to take risk index (std)	-0.162	-0.168	-0.162	-0.202
	(0.222)	(0.227)	(0.229)	(0.244)
Personality traits				
Openness (std)	-0.172	-0.258	-0.255	-0.129
	(0.194)	(0.188)	(0.191)	(0.210)
Conscientiousness (std)	0.243	0.297	0.287	0.145
	(0.213)	(0.214)	(0.216)	(0.236)
Extraversion (std)	0.219	0.286	0.280	0.213
	(0.179)	(0.180)	(0.182)	(0.212)
Agreeableness (std)	-0.248	-0.259	-0.250	-0.221
	(0.172)	(0.173)	(0.172)	(0.184)
Neuroticism (std)	0.114	0.155	0.147	0.091
	(0.206)	(0.213)	(0.214)	(0.216)
Demographics				
Male Dummy	-0.161	-0.166	-0.166	-0.243
	(0.271)	(0.272)	(0.273)	(0.279)
Age (std)	-0.157	-0.081	-0.081	-0.109
	(0.135)	(0.132)	(0.133)	(0.144)
N	188	178	178	171
χ^2	38.398**	33.560*	33.196*	30.049
Pseudo R^2	0.181	0.182	0.183	0.189
Average VIF predictors	1.902	1.900	1.904	1.938
NO INDIVIDUAL CONTROLS				
Monetary regret (std)	0.370***			
	(0.131)			
Effort regret (std)	-0.025			
	(0.122)			
w(0.33) (std)		-0.139		

	(0.113)		
w(0.67) (std)		-0.107	
		(0.124)	
Loss aversion index (std)			0.120
			(0.122)

This table reports the results from probit regressions with robust standard errors (in parentheses). Dependent variable: *Known Scheme Dummy*, which takes value one when a worker opted for the known scheme. *Individual controls* include ability, cognitive skills, personality and demographics. Target Treatment is a dummy that takes value one for Target treatments. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. **p*<0.10, ***p*<0.05, ****p*<0.01 and *****p*<0.001

Table A.8. Production, probability distortions and randomized incentives.

Treatments:	Random-Target Treatment		Random-Piece Rate Treatment	
Production (std)	[1]	[2]	[3]	[4]
Intercept	-0.667**	-0.674**	-0.325	-0.252
	(0.288)	(0.293)	(0.374)	(0.397)
Probability weights				
w(0.33) (std)	0.001		-0.243	
	(0.091)		(0.159)	
w(0.67) (std)		0.015		-0.094
		(0.056)		(0.178)
Regret motives				
Monetary regret (std)	0.364**	0.362**	0.228*	0.176
	(0.158)	(0.155)	(0.119)	(0.113)
Effort regret (std)	-0.038	-0.036	-0.041	-0.092
	(0.168)	(0.168)	(0.132)	(0.135)
Ability and cognitive skills				
Task ability (std)	0.474****	0.469****	0.426****	0.445****
	(0.134)	(0.133)	(0.129)	(0.135)
Raven (std)	0.046	0.042	-0.061	-0.063
	(0.119)	(0.124)	(0.118)	(0.129)
CRT (std)	0.011	0.020	0.240*	0.296**
	(0.165)	(0.167)	(0.144)	(0.137)
Risk attitudes				
Willingness to take risk index (std)	0.331***	0.330***	-0.119	-0.023
	(0.113)	(0.109)	(0.163)	(0.144)
Personality traits				
Openness (std)	0.259	0.264	-0.156	-0.198
	(0.233)	(0.234)	(0.136)	(0.149)
Conscientiousness (std)	-0.269	-0.266	0.112	0.196
	(0.238)	(0.238)	(0.229)	(0.232)
Extraversion (std)	-0.236	-0.237	-0.003	-0.030
	(0.156)	(0.156)	(0.132)	(0.142)
Agreeableness (std)	0.176	0.176	-0.001	0.058
	(0.127)	(0.127)	(0.211)	(0.226)
Neuroticism (std)	0.007	0.014	-0.084	-0.022
	(0.188)	(0.191)	(0.150)	(0.142)
Demographics				
Male Dummy	0.252	0.261	0.170	0.073
	(0.295)	(0.292)	(0.347)	(0.369)
Age (std)	-0.079	-0.079	-0.279**	-0.293**
	(0.095)	(0.095)	(0.114)	(0.117)

N	156	156	165	165
χ^2	88.020****	85.506****	144.018****	108.300****
R^2	0.423	0.422	0.500	0.463
Average VIF predictors	2.416	2.413	2.556	2.527
INTERACTION				
w(0.33) (std) × Task ability (std)	0.146		-0.076	
	(0.100)		(0.152)	
w(0.67) (std) × Task ability (std)		0.090		-0.144
		(0.092)		(0.257)
NO INDIVIDUAL CONTROLS				
w(0.33) (std)	0.047		-0.375***	
	(0.117)		(0.116)	
w(0.67) (std)		0.026		-0.142***
		(0.082)		(0.217)

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent Variable: *Production (std)*. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001

Table A.9. Production, regret and randomized incentives.

Treatments: *Production (std)*	Random-Target (Field Data) [1]	Random-Piece Rate (Field Data) [2]	Random-Target (Lab Data) [3]
Constant	-0.380	-0.244	0.291**
	(0.286)	(0.448)	(0.141)
<u>Loss aversion</u>			
Loss aversion index (std)	-0.099	-0.320	-0.104
	(0.123)	(0.252)	(0.088)
<u>Regret motives</u>			
Monetary regret (std)	0.465***	0.154	-
	(0.152)	(0.145)	
Effort regret (std)	-0.233*	0.099	-
	(0.138)	(0.201)	
<u>Ability and cognitive skills</u>			
Task ability (std)	0.438***	0.454***	0.423****
	(0.148)	(0.154)	(0.119)
Raven (std)	-0.021	-0.031	0.004
	(0.131)	(0.162)	(0.122)
CRT (std)	0.059	0.265*	0.114
	(0.189)	(0.157)	(0.157)
<u>Risk attitudes</u>			
Risk aversion index (std)	-	-	-0.189**
[number of safe choices, Holt & Laury, 2002]			(0.084)
Willingness to take risk index (std) [31]	0.396***	0.019	-
	(0.136)	(0.156)	
<u>Personality traits</u>			
Openness (std)	-0.038	-0.002	0.024
	(0.269)	(0.116)	(0.100)
Conscientiousness (std)	-0.309	0.098	-0.021
	(0.244)	(0.189)	(0.175)
Extraversion (std)	-0.011	-0.215	0.049
	(0.140)	(0.223)	(0.103)
Agreeableness (std)	0.298**	0.093	0.127
	(0.132)	(0.257)	(0.127)
Neuroticism (std)	0.015	-0.159	0.090
	(0.204)	(0.180)	(0.151)
<u>Demographics</u>			
Male Dummy	0.041	0.049	0.266
	(0.289)	(0.318)	(0.270)
Age (std)	0.081	-0.076	0.114
	(0.096)	(0.148)	(0.093)
N	162	148	52
Chi2	71.311****	255.893****	2.586**
R2	0.398	0.515	0.364
MeanVIF	2.594	2.744	1.545
NO INDIVIDUAL CONTROLS			
Loss aversion index (std)	-0.155	-0.340**	-0.025
	(0.108)	(0.164)	(0.085)

[31] We use the "Willingness to take risk index" instead of the "Risk-aversion index" in the field study because it was available for all workers.

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. *p<0.10, **p<0.05, ***p<0.01 and ****p<0.001

Table A.10. Production, curiosity and randomized incentives.

Treatments: *Production (std)*	Random-Target & Curiosity treatments [1]	Random-Piece Rate & Curiosity treatments [2]	All Random & Curiosity treatments [3]
Intercept	-0.007	-0.032	-0.091
	(0.145)	(0.164)	(0.112)
Curiosity Treatments	-0.135	-0.022	-0.076
	(0.111)	(0.106)	(0.074)
Piece Rate Treatments			0.079
			(0.074)
Ability and cognitive skills			
Task ability (std)	0.337****	0.309****	0.325****
	(0.064)	(0.080)	(0.051)
Raven (std)	0.074	0.035	0.070
	(0.070)	(0.081)	(0.053)
CRT (std)	0.153**	0.322****	0.226****
	(0.071)	(0.075)	(0.051)
Risk attitudes			
Willingness to take risk index (std)	0.156**	0.042	0.102**
	(0.063)	(0.063)	(0.046)
Personality traits			
Openness (std)	0.029	-0.090	-0.025
	(0.090)	(0.075)	(0.062)
Conscientiousness (std)	-0.004	0.063	0.033
	(0.095)	(0.098)	(0.073)
Extraversion (std)	-0.065	-0.070	-0.074
	(0.076)	(0.073)	(0.054)
Agreeableness (std)	0.037	0.075	0.056
	(0.092)	(0.067)	(0.055)
Neuroticism (std)	-0.054	0.007	-0.012
	(0.085)	(0.075)	(0.063)
Demographics			
Male Dummy	0.005	0.016	0.039
	(0.108)	(0.123)	(0.082)
Age (std)	-0.010	-0.075	-0.037
	(0.060)	(0.060)	(0.042)
N	520	523	1043
χ^2	104.798****	160.663****	216.565****
R^2	0.252	0.363	0.290
Average VIF predictors	1.788	1.788	1.687
NO INDIVIDUAL CONTROLS			
Curiosity Treatments	-0.132	-0.017	-0.073
	(0.122)	(0.131)	(0.089)

This table reports the results from linear panel regressions with robust standard errors (in parentheses) and random effects. Period fixed effects and scheduling fixed effects included. Dependent variable: *Production (std)*. Piece Rate (Curiosity) Treatments is a dummy variable that takes value one for the Piece Rate (Curiosity) treatments. The bottom panel displays coefficient estimates in the absence of individual controls. (std) refers to standardized variables. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001

C. Additional analyses

C.1. Target achievement

In this section, we investigate whether workers achieved their targets. In Figure C.1 we show the proportion of workers who achieved each of the three possible targets depending on their assigned target. For workers who were assigned a low target (13 in the Lab experiment and 50 in the Field experiment), the vast majority (98.1% and 74.8%) achieved the target. However, the share of workers who achieved the low target when assigned a medium (26 or 80) [high (39 or 110)] target, was significantly lower (81.6% and 64.6%; p-values = 0.004 and 0.078, proportion tests) [57.9% and 49.2%; p-values < 0.001]. This is probably due to a discouragement effect by which some workers, particularly those with low abilities, receiving medium or high targets may consider that they will not be able to achieve those targets and, as a result, quit producing. Interestingly, under random targets most of the workers did not quit and a remarkably similar proportion of workers achieved the low target (95.7% and 73.9% in the Lab and Field experiments), when compared to those who were assigned a low target in the Known-Target treatment (with respective p-values of 0.449 and 0.831). A similar pattern emerges for medium and high targets. The proportion of workers who achieved the medium [high] target in the Random-Target treatment (68.4% and 40.0%) [19.7% and 13.1%] was similar to the Known-Target treatment when workers were assigned the medium [high] target (62.3% and 37.8%; p-values = 0.440 and 0.689) [17.4% and 9.9%; p-values = 0.717 and 0.339]. In sum, randomized targets induced workers to achieve each of the potential targets as if the target value was known.

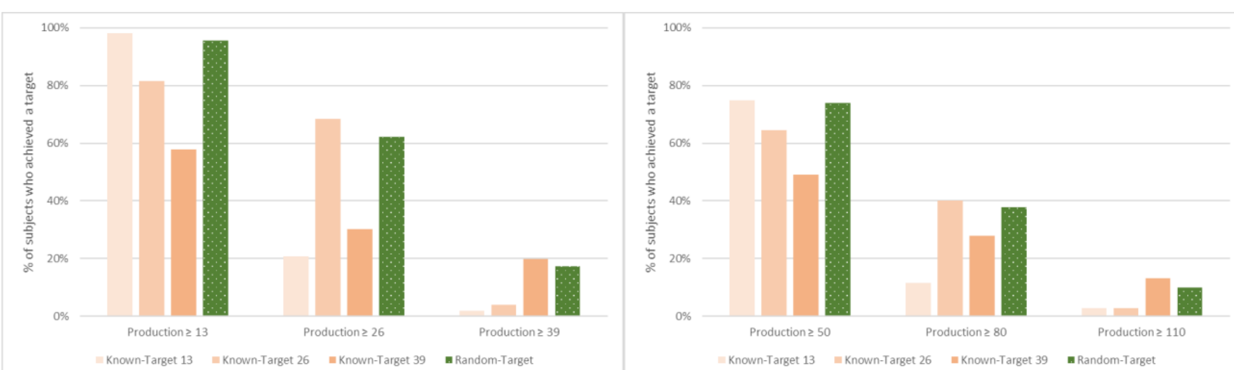


Figure C.1. Percentage of workers who achieved each target by treatment in the Lab experiment (left panel) and in the Field experiment (right panel).

Supplementary Material (online)

(Instructions)

Available at:

https://acesse.one/RandomizedIncentives-instructions

SM1. Lab Study

SM2. Field Study

SM2.1. Incentive treatments

SM2.2. Additional instructions