Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

Increasing worker motivation using a reward scheme with probabilistic elements

Adrian R. Camilleri [a,*], Katarina Dankova [b], Jose M. Ortiz [c], Ananta Neelim [d]

[a] University of Technology Sydney, Australia
[b] Central European Labour Studies Institute, Slovakia
[c] Zayed University, UAE
[d] Tasmanian Behavioural Lab, University of Tasmania, Australia

ARTICLE INFO

Keywords:
Experiment
Motivation
Performance
Reward
Incentives
Fairness
Justice
Ability

ABSTRACT

The purpose of this research was to investigate the effectiveness of a probabilistic reward scheme to motivate workers and increase their performance. Across seven experiments (three of which are in the online appendices) testing three different real effort tasks, we compared two novel probabilistic reward schemes with two traditional non-probabilistic reward schemes. In our flagship "single lottery" probabilistic scheme, worker performance was associated with the accumulation of lottery tickets in the worker's own personal lottery with a moderate jackpot on offer. It was possible for the worker to accumulate all tickets and thus guarantee the jackpot. We found that the single lottery scheme increased motivation and performance relative to other probabilistic and non-probabilistic schemes with the same expected values. There was also evidence that the single lottery scheme was particularly effective for lower-ability workers relative to the non-probabilistic schemes. We argue that the single lottery scheme uniquely benefited from optimism bias and the goal gradient effect. Considering perceptions of (un)fairness associated with probabilistic reward schemes – at least at first – we discuss what labor contexts are appropriate for the introduction of a probabilistic reward scheme.

1. Introduction

Understanding worker motivation is fundamental to the success of organizations (Kanfer & Chen, 2016). This is because the degree of worker motivation is associated with the quantity and quality of work, which in turn affects organizational performance (Bonner & Sprinkle, 2002). In line with such importance, labor makes up at least 70% of the average organizations' costs (Blinder, 2011).

The most common way for an organization to elicit worker motivation and performance is through financial compensation in the form of incentives and rewards. Although we think of "incentives" as financial compensation agreed to prior to work and "rewards" as financial compensation given after work, we proceed using the two terms interchangeably. According to agency theory (Eisenhardt, 1989), worker effort begins at zero and increases monotonically with the size of the compensation offered. Consistent with this prediction, several *meta*-analyses have confirmed that there is a positive association between financial compensation and worker performance (Condly et al., 2003; Garbers & Konradt, 2014; Jenkins et al., 1998; Weibel et al., 2009; Wiersma, 1992).

There are many kinds of financial compensation. These include forms of guaranteed compensation (e.g., salary), forms of variable compensation (e.g., bonuses), and forms of equity compensation (e.g., stock options) (Bonner et al., 2000). Of most interest for the current research is a form of variable compensation that we will refer to as performance-related pay. A performance-related pay incentive scheme is one in which at least some of a workers' financial compensation is based on their performance. Quite simply, if a worker produces more or performs better, then they get paid more (Lazear, 2000). Major theories of motivation such as expectancy theory (Van Eerde & Thierry, 1996; Vroom, 1964) and agency theory (Eisenhardt, 1989) predict that performance-related pay should increase worker output compared to flat wages, which empirical research supports (Cadsby et al., 2007; Lazear, 2000). It is, therefore, no surprise that the number of organizations using performance-related pay has increased over time (Lemieux et al., 2009; MacLeod & Parent, 1999; Mercer, 2018; PayScale, 2018).

This paper examines a novel performance-related pay incentive scheme that is probabilistic in nature. Under our flagship scheme – the

* Corresponding author.
E-mail address: adrian.camilleri@uts.edu.au (A.R. Camilleri).

https://doi.org/10.1016/j.obhdp.2023.104256
Received 30 January 2022; Received in revised form 3 May 2023; Accepted 18 May 2023

single lottery probabilistic scheme – worker performance is associated with the accumulation of lottery tickets in the worker's own personal lottery[1]. Winning the lottery earns a moderate jackpot. Across four main experiments (and three experiments reported in the online appendices), conducted both in the lab and online, and using three different tasks, we compared this single lottery probabilistic incentive scheme with other probabilistic and non-probabilistic incentive schemes with the same or very similar expected values. In Experiments 1 and 2, we found that the single lottery probabilistic incentive scheme produced the greatest motivation to work and the highest performance. In Experiments 3 and 4, we found two mechanisms driving the superiority of this particular implementation of a probabilistic reward scheme: optimism bias and the goal gradient effect. In Supplementary Study 3, we additionally implicated the timing of the risk uncertainty resolution as a contributing component. In Experiments 1, 2, 4, and Supplementary Study 2, we found tentative evidence that the single lottery probabilistic incentive scheme is more effective for low-ability workers. Finally, in Experiments 1 and 4, we found that the single lottery scheme is perceived by some as unfair but, in Supplementary Study 1, these concerns are ameliorated over time.

1.1. *The effectiveness of probabilistic rewards*

Probabilistic rewards are those that have uncertainty associated with the timing and/or the size of a reward. Probabilistic rewards are commonplace in a range of contexts. For example, many people engage in gambling – particularly in the form of lotteries – despite its negative expected value (Clotfelter et al., 1999). Some organizations and researchers have attempted to capitalize on people's enjoyment of gambling to encourage certain behaviors. For example, prize-linked savings accounts, which offer a chance to win a large prize as a function of the deposit amount, produce more savings than standard interest of the same expected value (Filiz-Ozbay et al., 2015; Guillén & Tschoegl, 2002). Probabilistic price promotions, such as a "1% chance it's free", increase the likelihood of and the number of purchases than comparable fixed price promotions (Lee et al., 2019; Mazar et al., 2017). Lottery prizes have been shown to increase adherence to medication and home-monitoring schedules (Kimmel et al., 2016; Kimmel et al., 2012; Sen et al., 2014; Volpp, Loewenstein, et al., 2008), physical activity (Patel et al., 2018), gym attendance (van der Swaluw et al., 2018a, 2018b), recycling (Diamond & Loewy, 1991; Luyben & Cummings, 1981; Maki et al., 2016), charity donations (Landry et al., 2006), and contributions to public goods (Corazzini et al., 2010).

Overall, these studies suggest that, in some circumstance, people enjoy probabilistic rewards and so these kinds of incentives can be strategically used to influence behavior. However, it is not straightforward to generalize these findings to performance-related worker pay due to some key differences. First, in financial and health contexts, probabilistic rewards are used to encourage people towards behaviors they are already motivated to perform whereas workers are primarily motivated by compensation because effort provision is costly. For that reason, probabilistic rewards are usually designed as a form of variable compensation on top of some form of guaranteed compensation. Second, in the financial and consumption contexts, the likelihood of winning the lottery is independent of the individual whereas in the labor context the likelihood of winning the lottery is related to the worker's effort and level of skill. The presence of effort requirements decreases the appetite for large uncertain rewards (Kivetz, 2003), underlines the potential for (un)fairness concerns, and highlights the value of evaluating a probabilistic reward scheme in light of workers' ability level.

In the labor context, there has been much less research on probabilistic rewards, and the research that does exist is inconclusive (Haisley, 2008; Miller et al., 2014; Pampino Jr et al., 2004; Wine et al., 2017). One early study offered a group of workers a lottery ticket contingent on performance associated with a 10% chance of earning 1, 2, or 5 state lottery tickets as well as entry into a weekly draw for $150 (Evans et al., 1988). Those given probabilistic rewards produced significantly more output than those allocated to an hourly pay system but significantly less output compared to those receiving a fixed bonus for meeting a specific target. Another study allocated lottery tickets to workers based on their relative performance to other workers (Cook & Dixon, 2006). The highest performer earned 3 tickets, the second-highest performer earned 2 tickets, and the third-highest performer earned 1 ticket. The weekly draw winner was awarded $50. Compared to baseline, performance was highest with the lottery reward system. Unfortunately, it is impossible to draw firm conclusions about the use of probabilistic rewards as a form of performance-related pay because the existing studies are under-powered (for example, $n = 2$ in Wine et al., 2017) and have low internal validity. For example, the studies do not compare probabilistic rewards with non-probabilistic rewards of the same expected value (e.g., Evans et al., 1988), and often combine multiple interventions non-systematically (e.g., feedback type in Cook & Dixon, 2006). In one of the better studies that avoided these issues, Haisley (2008) used what we call a "multiple lottery" scheme in which a new lottery (e.g., 1% chance of winning $10) was played after each task completion. She found that lottery-linked incentives did *not* lead to greater motivation. In sum, it remains unclear whether probabilistic rewards are motivating in a labor context and, if they are, when and why.

1.2. *Explanations for the effectiveness of probabilistic rewards*

The previous section outlined numerous contexts in which probabilistic rewards can be motivating and change people's behavior, sometimes to a greater degree than non-probabilistic rewards with the same expected value. From a theoretical standpoint, these findings are initially puzzling because normative and descriptive theories of risky choice predict that people are risk-averse for gains and prefer a reward of a certain magnitude over a reward of an uncertain magnitude (Kahneman & Tversky, 1979; Mineka & Hendersen, 1985; Von Neumann & Morgenstern, 1947). Moreover, some research suggests that people value an uncertain incentive even less than its lowest payoff (Gneezy et al., 2006; Simonsohn, 2009), although this effect has been qualified (Moon & Nelson, 2019; Rydval et al., 2009). In any case, we describe three explanations for why a probabilistic reward scheme could be motivating.

First, the uncertainty associated with lotteries increases excitement (Zuckerman, 2007) and, irrespective of the outcome, it is pleasant to resolve the uncertainty (Ruan et al., 2018). For example, in one study, members of a running club were given a reward based on the number of laps they ran (Shen et al., 2019). After each lap, half the participants received a certain 5 points whereas the other half received either 5 or 3 points with equal probability. Points were later converted into a gift card at a local cafe. On average, those allocated to the uncertain reward scheme ran more laps. Evidently, a probabilistic reward scheme benefits from both the positive utility of uncertainty resolution as well as the positive utility of the reward itself. There is even evidence that receiving unpredictable (vs. predicable) rewards is correlated with activation of dopamine projection sites in the brain, which may indicate a relatively higher degree of pleasure (Berns et al., 2001).

Second, people display optimism bias in that they believe that they are less likely than others to experience negative events and more likely than others to experience positive events (Helweg-Larsen & Shepperd, 2001). For example, in one study, students rated their likelihood of experiencing a range of positive events – including liking their post-graduation job and living past the age of 80 – as significantly more likely compared to other student's chances (Weinstein, 1980). Such

[1] Note that in all probabilistic reward schemes considered here, the worker is entered into their own personal lottery and are competing only with chance – and not other workers – for the reward. We discuss "tournament" style incentive reward schemes in the General Discussion.

overoptimism has also been found using tasks associated with the occurrence of aleatory events (Irwin, 1953). For example, Seybert and Bloomfield (2009) presented participants with a deck of cards containing a mix of favorable and unfavorable cards. On repeated trials, participants had to estimate the probability of a favorable card being drawn and then bet on that card being drawn. Behavior was consistent with "wishful betting" such that desirable cards were predicted as more likely than their relative frequency.

Third, people tend to treat probabilities non-linearly. Prospect theory proposes an inverse S-shaped probability weighting function implying that relatively small probabilities are overweighted and relatively large probabilities are underweighted (Kahneman & Tversky, 1979). Such a propensity would increase the attractiveness of a lottery with a small chance of a big reward (although, this could be countered by a sufficiently concave value function). For example, most people prefer a 5% chance to win $10,000 over a 100% chance to obtain $500 even though both have the same expected value; this is often termed the "possibility effect" (Wu & Gonzalez, 1996).

Considering these theories and the literature review, we developed a general hypothesis that a probabilistic reward scheme would be superior to a non-probabilistic reward scheme. We sought to evaluate reward schemes on two dimensions: motivation and performance (Bonner & Sprinkle, 2002). Motivation includes the worker's decision of what activity to engage in, as well as how long and how intensely to engage in that activity. Motivation is often measured by absenteeism and task choice (Kanfer, 1990). Performance includes both work quantity as well as work quality. Performance is often measured by the number of outputs per unit of time as well as the level of workmanship relative to what is standard or expected.

H1: Probabilistic rewards (vs. non-probabilistic rewards with the same expected value) will (i) increase motivation to work and (ii) increase work performance.

1.3. Goals

Broadly speaking, a probabilistic rewards scheme can be conceptualized in one of two different ways depending on whether additional worker output changes the number of lotteries participated in or changes the probability of winning the lottery. The first way, which we call the "multiple lottery" probabilistic reward scheme, rewards additional work by triggering a lottery at specific output thresholds. For example, for every 10 widgets produced, a worker has a 20% chance of winning a $100 bonus payment. Consequently, a worker who produces 20 widgets would play two lotteries, each with a 20% chance of winning $100, which has an expected value of $40 (i.e., $20 for each of the two lotteries). The second way, which we call the "single lottery" probabilistic reward scheme, rewards additional work through the accumulation of tickets in a single, pre-scheduled lottery. For example, for every 10 widgets produced, an additional ticket is earned in a lottery comprising five total tickets to win a $100 bonus payment. Consequently, a worker who produces 20 widgets would play one lottery with a 40% chance of winning $100, which has an expected value of $40. Note that, under the multiple lottery scheme, the chance of winning the bonus is fixed and additional worker output increases the number of lotteries the worker participates in whereas, under the single lottery scheme, participants only play one lottery and additional worker output increases the probability of winning. Under the second scheme, with extraordinary performance, it may even be possible to increase the chance of winning the single lottery to 100%. We believe this difference is crucial for the optimal design of a probabilistic rewards scheme because the second scheme better taps into an additional source of motivation: goals. However, to our knowledge, the second design has never been tested.

A large literature of research demonstrates that the pursuit of specific, difficult goals leads to higher performance compared to the pursuit of non-specific or easy goals (Locke & Latham, 2002). This relation between goals and performance is strongest when people are committed to the goal (Klein et al., 1999). According to goal setting theory, financial incentives are one way in which commitment to difficult goals can be enhanced (Locke et al., 1988; Wright, 1992). As such, there is a positive relationship between goal-contingent financial incentives and performance (Corgnet et al., 2015; Locke & Latham, 1990).

Making progress towards a goal is motivating (Bonezzi et al., 2011; Huang & Zhang, 2011; Soman & Shi, 2003; Wallace & Etkin, 2018). For example, consumers are more motivated when they have progressed 2 steps out of the required 10 than 0 steps out of the required 8, presumably because the former situation indicates that some progress has been made to the goal (Nunes & Dreze, 2006). Not only is goal progress motivating but it is increasingly more motivating as the completion of a specific goal nears – this is called the goal gradient effect (Hull, 1932). For example, consumers in a café loyalty reward program tend to purchase more coffee the closer they are to earning a free coffee (Kivetz et al., 2006). An explanation for the goal gradient effect is that a specific goal becomes a salient reference point that divides the possible outcome space into "gain" and "loss" and, reminiscent of prospect theory's value function (Kahneman & Tversky, 1979), the value associated with goal progress is steeper closer to the reference point (Heath et al., 1999). Consequently, each unit of marginal progress towards the goal is associated with increasingly larger amounts of utility.

We argue that the single lottery – compared to the multiple lottery – probabilistic reward scheme benefits more from goal progress and the goal gradient effect. To illustrate, consider Table 1 and the details of the two probabilistic reward scheme examples mentioned earlier. Under both schemes, the goal is to get the $100 reward. Under the single lottery scheme, the way to achieve the goal is specific and unequivocal: by completing 50 widgets, thus accumulating all lottery tickets, thus guaranteeing a lottery win. Under this scheme, 50 is a strong reference point, additional worker output corresponds directly with goal progress, and there is a relatively low amount of uncertainty along the way. In contrast, under the multiple lottery scheme, the way to achieve the goal is less specific and equivocal: it is unclear how many widgets need to be completed to achieve the goal; it could be as little as 10 or perhaps more than 50. As a result, under this scheme, 10, 20, 30, 40, and 50 are all

Table 1
Hypothetical Potential Outcomes Under Different Reward Schemes and Performance.

Widgets Completed	Outcome			
	Probabilistic Reward Schemes		Non-Probabilistic Reward Schemes	
	Single Lottery	Multiple Lottery	Piece Rate	Lump Sum
0	1 attempt at 0% chance of $100	0 attempts at 20% chance of $100	Bonus of $0	Bonus of $0
10	1 attempt at 20% chance of $100	1 attempt at 20% chance of $100	Bonus of $20	Bonus of $0
20	1 attempt at 40% chance of $100	2 attempts at 20% chance of $100	Bonus of $40	Bonus of $0
30	1 attempt at 60% chance of $100	3 attempts at 20% chance of $100	Bonus of $60	Bonus of $0
40	1 attempt at 80% chance of $100	4 attempts at 20% chance of $100	Bonus of $80	Bonus of $0
50	1 attempt at 100% chance of $100	5 attempts at 20% chance of $100	Bonus of $100	Bonus of $500

Note: The expected value of the bonus is the same for equal levels of performance under the Single Lottery, Multiple Lottery, and Piece Rate schemes. For example, under the Piece Rate scheme, for completing 20 widgets the worker will get $40. Under the Single Lottery scheme, the worker will play one lottery with 40% chance of winning $100, i.e., an expected value of $40. Under the Multiple Lottery scheme, the worker will play two lotteries, each with a 20% chance of winning $100, i.e., an expected value of $40 ($20 for each of the two lotteries).

weak reference points, additional worker output corresponds only indirectly with goal progress, and there is a relatively high amount of uncertainty along the way. The multiple lottery scheme is also associated with repeated opportunities for either satisfaction if the lottery is won or, more likely, dissatisfaction if the lottery is lost. Repeated failure to achieve the goal can produce negative emotions and undermine motivation (O'hora & Maglieri, 2006).

H2: A single lottery (vs. multiple lottery with the same expected value) probabilistic reward scheme will (i) increase motivation to work and (ii) increase work performance.

1.4. Ability

A potential boundary condition of the probabilistic reward effect (H1) is worker ability. A worker's knowledge, skills, and aptitude – what we will collectively refer to here as "ability" – has always been at the center of research aimed at understanding worker motivation and performance (Sackett et al., 2017). This is because research demonstrates that ability is directly related to performance (Schmidt et al., 1986), particularly for tasks requiring judgment and decision-making or where intrinsic motivation is relatively low (Camerer & Hogarth, 1999). There is, however, a paucity of research investigating how ability and incentive reward schemes interact.

One of the major criticisms of performance-related pay reward schemes is that they can be demotivating to most workers at the expense of a few high performers (Kennedy, 1995; Meyer, 1975; Suff et al., 2007). For example, consider the two popular non-probabilistic reward schemes displayed in Table 1 – the "piece rate" and "lump sum" schemes (DellaVigna & Pope, 2018; Zhang & Gao, 2016) – and contrast the prospects of a low-ability worker and a high-ability worker. The low-ability worker may be discouraged by both non-probabilistic reward schemes because their potential bonus is so low compared to that of a high-ability worker. We argue that a probabilistic reward scheme may help avoid undermining the motivation of low-ability workers for two reasons. First, probabilistic reward schemes give low-ability workers the opportunity to earn the same bonus as high-ability workers. Second, given that the single lottery scheme determines the lottery outcome only at the end of the work period, it allows optimism to drive motivation during the entire period. Therefore, we hypothesized that worker ability is a moderator of the probabilistic rewards effect (i.e., of H1). It is likely that this anticipated moderation is driven by perceived worker ability, at least initially, and that perception likely becomes increasingly more positively correlated with objective ability as the worker observes their performance relative to initial expectations and the performance of other workers.

H3: Probabilistic rewards (vs. non-probabilistic rewards with the same expected value) will produce (i) higher motivation to work and (ii) higher work performance for those with relatively low- (vs. high-) ability.

1.5. Fairness

The anticipated motivating effect of probabilistic rewards may be undermined by perceptions of (un)fairness. Perceptions of fairness predict a range of important worker attitudes and behaviors, including performance (Cohen-Charash & Spector, 2001). Two of the most important forms of perceived fairness are distributive justice and procedural justice (Colquitt, 2001): Distributive justice relates to perceptions of how fairly outcomes were allocated (Deutsch, 1975) whereas procedural justice relates to perceptions of how fair the process was that determined how the outcomes were allocated (Cappelen et al., 2007; Lind & Tyler, 1988).

One criticism of performance-related pay schemes is that they can create dissatisfaction if they are perceived to be unfair (Isaac, 2001). Probabilistic reward schemes may exacerbate this fairness concern. This is because under a single lottery reward scheme, it is possible for a low-ability worker who wins the lottery to earn a larger reward than a higher-ability worker who loses the lottery. Such a result could harm perceptions of distributive justice because rewards are not entirely based on innate abilities and working hard, and also not proportional to performance, thus violating equity theory (Walster et al., 1978).

It is less clear whether negative perceptions will extend to procedural justice. Some have argued that random mechanisms are "the embodiment of fair allocation" (Oberholzer-Gee et al., 1997, p. 89) because allocation cannot be based on favoritism or personal characteristics. Nevertheless, surveys comparing the fairest ways to allocate scarce resources find that people judge random procedures to be unfair (Raux et al., 2009; Savage & Torgler, 2010). However, other surveys have revealed that the use of lotteries to allocate scarce resources is considered fairer when those entered into the lottery are first pre-screened for merit (Brickman et al., 1981).

H4: Probabilistic rewards (vs. non-probabilistic rewards with the same expected value) will be perceived as less fair, particularly in terms of distributive justice.

1.6. Overview of studies

In what follows, we report four experiments. In Experiment 1, using an incentive-compatible real effort task requiring participants to quickly move sliders, we demonstrate that work performance is higher under a single lottery (SL) scheme than a multiple lottery scheme, piece rate scheme, and lump sum scheme. In Experiment 2, using an incentive-compatible real effort task requiring participants to quickly solve math puzzles, we replicate the effect and extend it to a pure motivation measure. In Experiment 3, we demonstrate that the SL scheme benefits from optimism bias. In Experiment 4, we explore to what extent the superiority of the SL scheme is driven by the fact that it is possible to guarantee the reward. In Experiments 1, 2, and 4 we also explore to what extent the superiority of the SL scheme is affected by the relative skill of the workers. In Experiments 1 and 4, we also explore perceptions of reward scheme fairness and effectiveness.

In addition to the four main experiments, we also report three additional lab experiments in the Online Appendices. In Supplementary Study 1, we investigate how SL scheme perceptions of unfairness diminish once workers recognize that there is a strong positive correlation between performance and rewards after multiple periods. In Supplementary Study 2, we again investigate the importance of ability as a moderator this time using a more valid measure of ability. In Supplementary Study 3, using an incentive-compatible real effort task requiring participants to quickly count letters in a sentence, we examine the extent to which the SL (vs. ML) scheme is superior due to it resolving the uncertainty only at the end of the period.

2. Experiment 1

The purpose of Experiment 1 was to test H1ii, H2ii, and H3ii. To do this, we asked each participant to complete a real-effort task under different incentive schemes where bonus amounts and performance targets were varied. We also asked some follow-up questions to begin our exploration of participants' perceptions of the different schemes and thus test H4.

2.1. Methods

The experiment was carried out at a laboratory for conducting economic experiments in a major university in England. The 115 participants (44% male) were recruited from the student database based on availability and budget. Each participant received a show-up fee of GBP £3.00.

An overview of the procedure is presented in Fig. 1. The study comprised of a real effort task followed by a questionnaire. The real effort task was the "slider task" used by Gill and Prowse (2012), which

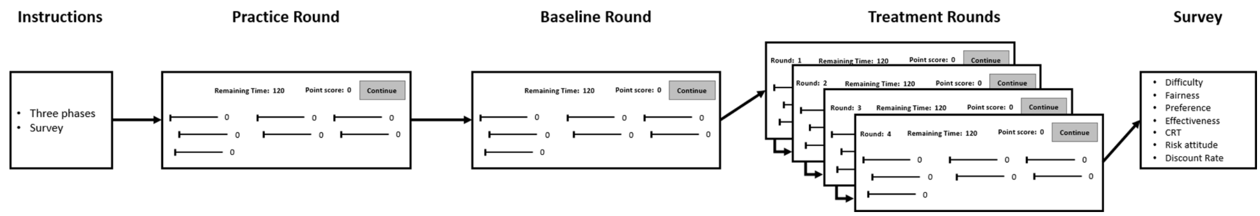

Fig. 1. An Overview of the Procedure Used in Experiment 1.

was selected because it is relatively straight-forward to perform without training. In the slider task, the participant was presented with a screen consisting of seven sliders. Each slider had a handle that could occupy the range of possible positions between 0 and 100. The participant's task was to move the handle of as many sliders as possible to the midpoint position (i.e., 50) in 120 seconds. The initial position of each handle was at 0. Once each handle on the screen had been moved to the midpoint of its respective slider, a button appeared that allowed the participant to proceed to the next screen where the next set of seven sliders appeared. This continued until all sliders were completed or time ran out.

In the first phase of the slider task, which we refer to as the Practice Round, the participant was presented with a total of 42 sliders. This phase was not incentivized and allowed the participant to become familiar with the task. In the second phase, which we refer to as the Baseline Round, the participant was presented with a total of 42 possible sliders across six screens. This phase, which provided our measure of natural ability, was incentivized with a flat rate of GBP £0.1 per slider competed. The third phase, which we refer to as the Treatment Rounds, was the same as the Baseline Round except that it occurred four times – each time under a different bonus incentive scheme. For each of these Treatment Rounds, the participant was allocated a new 120 seconds. The order of the incentive scheme was random. Note that the bonus incentive was paid in addition to the GBP £0.1 piece rate. The four different bonus incentive schemes were:

- **Lump Sum (LS)**: If 28 sliders are completed, earn a bonus of GBP £27.00.
- **Piece Rate (PR)**: For every 7 sliders completed, earn a bonus of GBP £2.50, with a maximum bonus of GBP £10.
- **Single Lottery (SL)**: For every 7 sliders completed, accrue an additional 25% chance of winning a lottery worth GBP £10.
- **Multiple Lottery (ML)**: For every 7 sliders completed, play a lottery with a 25% chance to win a bonus of GBP £10 with a maximum of four lottery attempts. The lottery was played immediately upon completing 7 sliders. The countdown paused while the lottery was carried out.

Note that the expected value for the piece rate, single lottery, and multiple lottery incentive schemes was identical. We attempted to also match the expected value of the lump sum scheme by setting the bonus amount considering the performance observed in the original Gill and Prowse (2011) data.

In the follow-up questionnaire, we measured the participant's perceived difficulty to earn the bonus (1 = "Extremely easy"; 7 = "Extremely difficult"), perceived fairness of the bonus scheme (1 = "Extremely unfair"; 7 = "Extremely fair"), strength of preference for each bonus scheme if playing an additional round (1 = "Extremely weak"; 7 = "Extremely strong"), and perceived effectiveness of each bonus scheme if the participant was a manager tasked with motivating employees (1 = "Ineffective"; 7 = "Effective"). In addition, we measured the participant's cognitive reflection tendency (Frederick, 2005), risk attitude (Dohmen et al., 2011), psychological discount rate (Kirby & Maraković, 1996), and demographics. At the end of the experiment, the participant was thanked and paid with their show-up fee, their earnings from the Baseline Round, and their earnings during one randomly selected Treatment Round. The average time to complete the experiment was approximately 40 minutes.

2.2. Results

In the Practice Round, participants completed a mean of 14.4 sliders. In the Baseline Round, participants completed a mean of 19.5 sliders, which was significantly more than in the Practice Round ($p < 0.001$). When the bonus schemes were introduced, the mean number of sliders completed successively increased further to a ceiling level.[2] Therefore, to better evaluate the effect of the different incentive schemes, we restricted our analysis to the first scheme encountered by each participant during the Treatment Rounds. This effectively converted our design and analysis approach from within-subjects to between-subjects. However, for the purposes of completeness, we provide analysis using all rounds in Appendix A.

Fig. 2 displays the mean number of sliders completed under the different incentive schemes. To compare the schemes, we used OLS regression where the dependent variable was the number of slider tasks completed under the first scheme encountered by each participant during the Treatment Rounds and the independent variables were indicator variables for each of the various schemes. Using the estimates of the scheme coefficients from the regression, we constructed scheme differences in mean performance (see Table 2). Supporting H1ii, we found that the participants completed 6.3 and 9.5 more slider tasks in the SL scheme relative to the PR and LS schemes, respectively. By contrast, the ML scheme produced only 1.5 and 4.6 more slider completions compared to the PR and LS schemes, respectively. Supporting H2ii, participants in SL scheme completed more tasks than those in the ML scheme. In the *Heterogenous Effects Across Ability* section, we discuss the impact of ability. Further, in the *Fairness and Effectiveness Perceptions* section, we discuss fairness perceptions and preference for a reward scheme. The other measures (e.g., cognitive reflection, risk attitude, demographic factors) were exploratory and we do not present results related to them.

2.3. Discussion

The results of Experiment 1 support our primary hypothesis that a probabilistic reward scheme can be more motivating than a non-probabilistic reward scheme with the same expected value. Specifically, participants completed significantly more tasks when their performance-based bonus was contingent on a lottery than when the bonus was more certain. Our follow-up analyses revealed that the incentive scheme that was primarily driving this effect was the single lottery scheme. Under this scheme, participants' performance was

[2] The mean number of slider completions was 32.1 for LS, 30.9 for PR, 31.8 for SL, and 31.0 for ML across the Treatment Rounds. None of the differences were statistically significant. We believe this is a ceiling effect given that the average number of correct completions was greater than the set performance targets. Furthermore, we find systematic evidence of learning over time: between first two blocks of the third phase (28.1 vs 31.1, p-value < 0.01, t-test) and second and third blocks of the third phase (31.1 vs 32.9, p-value < 0.01, t-test).

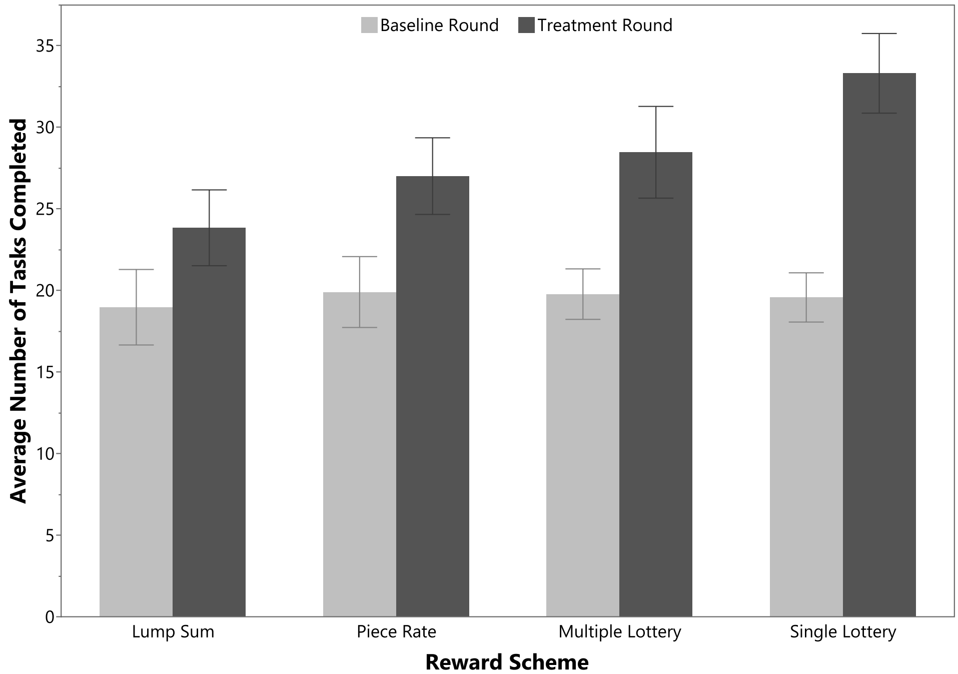

Fig. 2. Average Number of Tasks Completed During Baseline Round and the First Treatment Round Split by Incentive Scheme in Experiment 1 Error bars represent 95% confidence intervals.

Table 2
Motivation and Performance Results from Experiment 1, Experiment 2, and Experiment 4.

Differences	Experiment 1				Experiment 2				Experiment 4	
	Performance		Motivation		Performance		Motivation		Performance	
	Diff.	*p-val.*	Diff.	*p-val.*	Diff.	*p-val.*	Diff.	*p-val.*	Diff.	*p-val.*
SL vs PR	6.3	0.000*	0.18	0.000*	0.19	0.000*	0.07	0.010*	0.07	0.006*
SL vs LS	9.5	0.000*	0.12	0.007*	0.11	0.038				
SL vs ML	4.8	0.006*	0.09	0.034	0.12	0.020				
SL vs SL + U							0.08	0.012*	0.08	0.004*
ML vs PR	1.5	0.404	0.09	0.079	0.07	0.167				
ML vs LS	4.6	0.009*	0.03	0.542	−0.01	0.825				
SL + U vs PR							−0.00	0.946	−0.01	0.769

LS = Lump sum, PR = Piece Rate, SL = Single lottery, ML = Multiple lottery, SL + U = Single lottery + Uncertainty. Mean differences are based on regression coefficients estimates for schemes. For Experiment 1, the value for mean difference indicates the difference in number of sliders completed between the relevant groups. For example, the mean difference of 6.3 between SL and PR implies 6.3 more sliders were completed under the SL scheme relative to PR scheme. For Experiment 2 and 4, the value for mean difference implies the difference in probability of choosing (or correctly completing) the task across the relevant group. For example, the mean difference of 0.18 in motivation between SL and PR implies there was an 18% higher chance of participants choosing to do that math task under the SL scheme relative to the PR scheme. *p*-values are obtained from post-hoc tests on regression estimates. * denotes that the *p*-value is statistically significant after Bonferroni-Holm correction for multiple hypothesis testing. Five hypotheses were tested in Experiment 1 and 2. Two hypotheses were tested in Experiment 4.

highest when additional work moved them closer to accumulating all tickets in their own private lottery.

3. Experiment 2

The purpose of Experiment 2 was to examine the efficacy of the probabilistic incentive schemes in a different labor market context. In this experiment, we moved to a more cognitive real-effort task. One reason for this change is that different types of jobs are often associated with specific types of incentive schemes (MacLeod & Parent, 1999). Moreover, research suggests that the effectiveness of an incentive can vary depending on the type of work (Bonner & Sprinkle, 2002; Garbers & Konradt, 2014). Therefore, the real-effort task employed in this experiment was cognitive in nature. The new task also allowed us to distinguish between motivation and performance, and thus test both elements of H1, H2, and H3.

3.1. *Methods*

The experiment was carried out online. The 251 participants (39% male; M_{age} = 33.2 years) were recruited from Amazon's Mechanical Turk (AMT). The sample size was chosen based on an expected medium-to-large effect size, 90% power, and alpha of 0.05, which yielded a minimum group size of 50.[3] Each participant received a show-up fee of US$3.63.

[3] The effect size of comparisons of the SL scheme to the other schemes in Experiment 1 was large (Cohen's d = 1.48, 0.99, 0.71 for SL vs. LS, SL vs. PR, and SL vs. ML, respectively). Another way of calculating the required sample size would have been to use a framework that considers the repeated nature of the data. Our sample size of 250 participants, each making 30 decisions (correlation of 0.9 within decisions) would be able to detect a medium effect with 0.92 probability under this framework.

The study comprised of a real effort task followed by a questionnaire. The real effort task was the "math task" used by Goswami and Urminsky (2017). The math task contained 30 rounds. The number of rounds was not disclosed to participants. At the beginning of each round, the participant was asked to decide what to do during that round: to work on a math task for 30-seconds or to watch a funny advertisement video for 30-seconds. The decision to attempt the math task was our operalization of motivation. The math task itself consisted of searching a 3×4 grid containing twelve numbers and correctly selecting the unique two number combination that summed to 10 (see Appendix B). The number of math tasks completed was our operalization of performance.

The participant could earn additional money depending on their performance and the incentive reward scheme they were allocated to. The different schemes were:

- **Control (C):** No additional incentives are available.
- **Lump sum (LS)**: If 24 math tasks are completed, earn a bonus of $1.04.
- **Piece Rate (PR)**: For every 3 math tasks completed, earn a bonus of $0.02[4], with a maximum bonus of $0.16.
- **Single Lottery (SL)**: For every 3 math tasks completed, accrue an additional 12.5% chance of winning a lottery worth $0.16.
- **Multiple Lottery (ML)**: For each 3 math tasks completed, play a lottery with a 12.5% chance to win a bonus of $0.16 with a maximum of eight lottery attempts. The lottery was played immediately upon completing 3 math tasks.

As in Experiment 1, the expected value for the piece rate, single lottery, and multiple lottery incentive schemes were identical. We also attempted to match the expected value of the lump sum scheme by setting the bonus amount based on the performance of workers in the control condition, which was completed first. In our calculations, we assumed that the introduction of a bonus would increase performance by 15%. Using the extrapolated performance, we found that the expected pay-off from the piece rate, single lottery, and multiple lottery would be $2.1 and the number of people in the lump sum scheme that would be eligible for the bonus to be 2, leading to a bonus amount of $1.05 for each, which was reduced to $1.04 to ensure that the bonus amount in the Lump sum scheme was exactly 6.5 times that of the piece rate, single lottery, and multiple lottery incentive schemes.

In the follow-up questionnaire, we measured the participant's risk preference (Dohmen et al., 2011), numeracy (Cokely et al., 2012), cognitive reflection tendency (Frederick, 2005; Primi et al., 2016), and demographics. Given the mathematical nature of the task, participant's score on the dynamic version of the Berlin Numeracy Test, which ranged between 1 and 4, was the basis of our ability variable. At the end of the experiment, the participant was thanked and then paid with their show-up fee and their bonus earnings. The average time to complete the experiment was approximately 35 minutes.

3.2. Results

With respect to motivation, averaging across groups and rounds, participants attempted the math task 76.4% of the time. However, as shown in Fig. 3, motivation depended on the incentive scheme. To analyze these data, we used a random-effects Probit regression (random intercepts) to account for multiple observations per participant. The dependent variable was the attempt of math task (0 = *No* and 1 = *Yes*) and the independent variables were the different incentive schemes. In Table 2, we present the estimated means at the margin for each group and a post-estimation test examining mean differences across groups. Supporting H1i, those incentivized under the SL scheme were more likely to attempt the math task than those in the PR and LS schemes. However, those in the ML scheme were no more likely to attempt the math task than those in the PR and LS schemes. Supporting H2i, those incentivized under the SL scheme were more likely to attempt the math task than those in the ML scheme.

With respect to performance, averaging across groups and rounds, participants completed 12.5 out of 30 math tasks, which corresponds approximately to a 42% completion rate. To analyze these data, we used a panel regression framework described above, where task completion was the dependent variable (0 = *Not completed* and 1 = *Completed*) and the independent variables were the different incentive schemes. Table 2 presents the mean differences across groups. We again observed similar patterns. Supporting H1ii, performance under the SL scheme was superior to that of the PR and LS schemes. However, the ML scheme did not produce better performance relative to the PR and LS schemes. Supporting H2ii, those incentivized with the SL scheme completed more math tasks than those with the ML scheme. We discuss the effects of ability in the *Heterogenous Effects Across Ability* section.

3.3. Discussion

The results of Experiment 2 provide further support for the relative superiority of a probabilistic rewards scheme. In a more cognitive task relative to that used in Experiment 1, we found that participants attempted and successfully completed more tasks when the performance-based pay was contingent on a single lottery. Under this scheme, both participants' motivation and performance were the highest and significantly better than that in the other schemes, including the multiple lottery reward scheme. In the following two experiments, we attempted to better understand the causal mechanisms underlying the probabilistic incentive schemes, particularly the superiority of the single lottery scheme.

4. Experiment 3

The purpose of pre-registered Experiment 3 was to shed light on a potential mechanism driving the effectiveness of probabilistic reward schemes. According to our theorizing, probabilistic reward schemes benefit from *optimism bias*; that is, the tendency to believe that one has a better chance of good outcomes than is objectively true. In the context of the single lottery probabilistic reward scheme, this optimism bias could manifest at a group level as higher expected rates of winning the bonus than is statistically possible. For example, those who expect to complete 12 math tasks have a 50% chance of earning the bonus and a 50% chance of earning no bonus. An optimism bias would be apparent if more than 50% of those who expect to complete 12 math tasks believe they will win the lottery and earn the bonus. To test this aspect of the theoretical mechanism, we explained either the piece rate or single lottery reward scheme to participants and then asked about their expectations prior to beginning the task.

4.1. Methods

The experiment was carried out online. The 263 participants (62.6% male; M_{age} = 40.5 years) were recruited from AMT. Each participant who completed the study received US$1.26.

The study used the same real effort task as in Experiment 2 followed by a questionnaire. Participants were randomly allocated to one of two bonus reward schemes: (1) Piece rate (PR) where, for every 3 math tasks completed, a bonus of $0.02 was earned, with a maximum bonus of $0.16; and (2) Single Lottery (SL) where, for every 3 math tasks completed, an additional 12.5% chance of winning a lottery worth $0.16 was accrued. Following our pre-registration plan, we oversampled the SL group because of our intention to explore its interaction with ability.

After the instructions and comprehension questions, participants were asked three questions about their expectations. First, we asked

[4] The bonus amounts were selected after conducting pilot studies to avoid ceiling and floor effects.

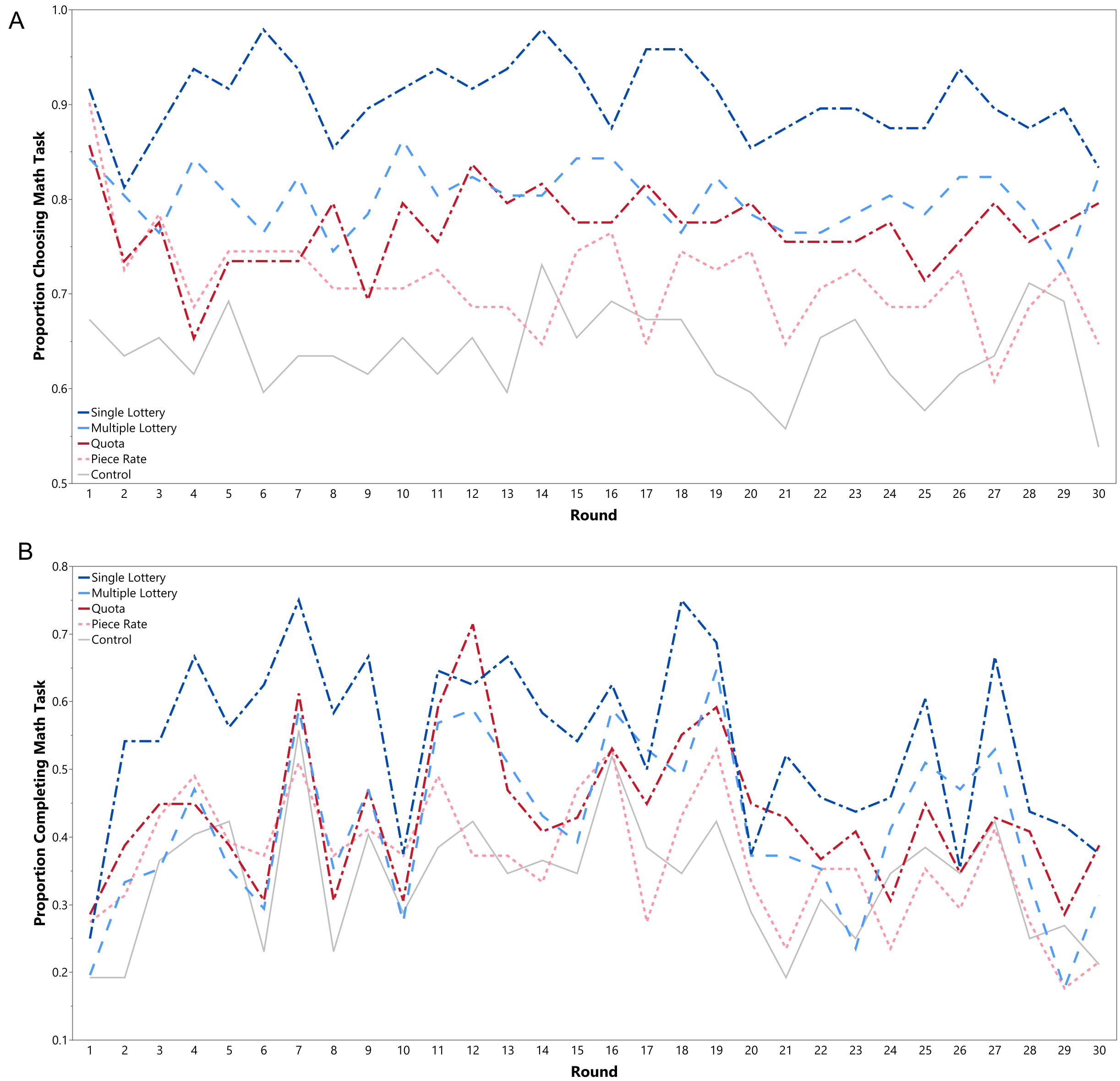

Fig. 3. (A) Proportion Choosing to Complete the Math Task and (B) Proportion Completing the Math Task Split by Round and Incentive Scheme in Experiment 2.

about their expected number of math task attempts (ranging between 0 and 30). Second, we asked about their expected number of math task completions (ranging between 0 and their expected number of math task attempts). Finally, we asked about their expected bonus amount. For those in the PR group, this final question was redundant considering their response to the second question and, indeed, the only available option presented on screen was the bonus amount that would be received considering the expected number of math task completions (e.g., for 3-to-5 expected task completions, $0.02 was the only option available). For those in the SL group, this final question was presented as a binary choice between $0.00 (i.e., losing the lottery and earning no bonus) and $0.16 (i.e., winning the lottery and earning the maximum bonus). After completing the expectation questions, the participants were told that they did not need to complete the task and could move to the next part of the study. The follow-up questionnaire was the same as that used in Experiment 2. The average time to complete the experiment was approximately 10 minutes.

4.2. Results

The average expected number of tasks attempted and completed was 10.6 and 9.3, respectively (see Fig. 4 for the data split by reward scheme). Following our pre-registered analysis plan, we conducted two t-tests, which revealed no significant differences between reward scheme in expected number of tasks attempted ($p = 0.844$) or completed ($p = 0.911$).

Next, we turned to the degree of optimism (or pessimism) exhibited by participants in the SL group. In Fig. 5, the *x*-axis reports the objective chance of winning the lottery based on the participant's expected number of task completions. The *y*-axis reports the participant's expected chance of winning the bonus based on the participant's expected bonus payment. Note that participants could only indicate $0 or $0.16 as the bonus they would receive. These responses are represented by the black dots along the 0% and 100% chance of winning horizontal lines. Following our pre-registered analysis plan, we generated an "optimism score" for each participant in two steps. First, we ran a Probit regression where the expected bonus ($0 or $0.16) was the dependent variable. The independent variable was the objective probability of winning the bonus

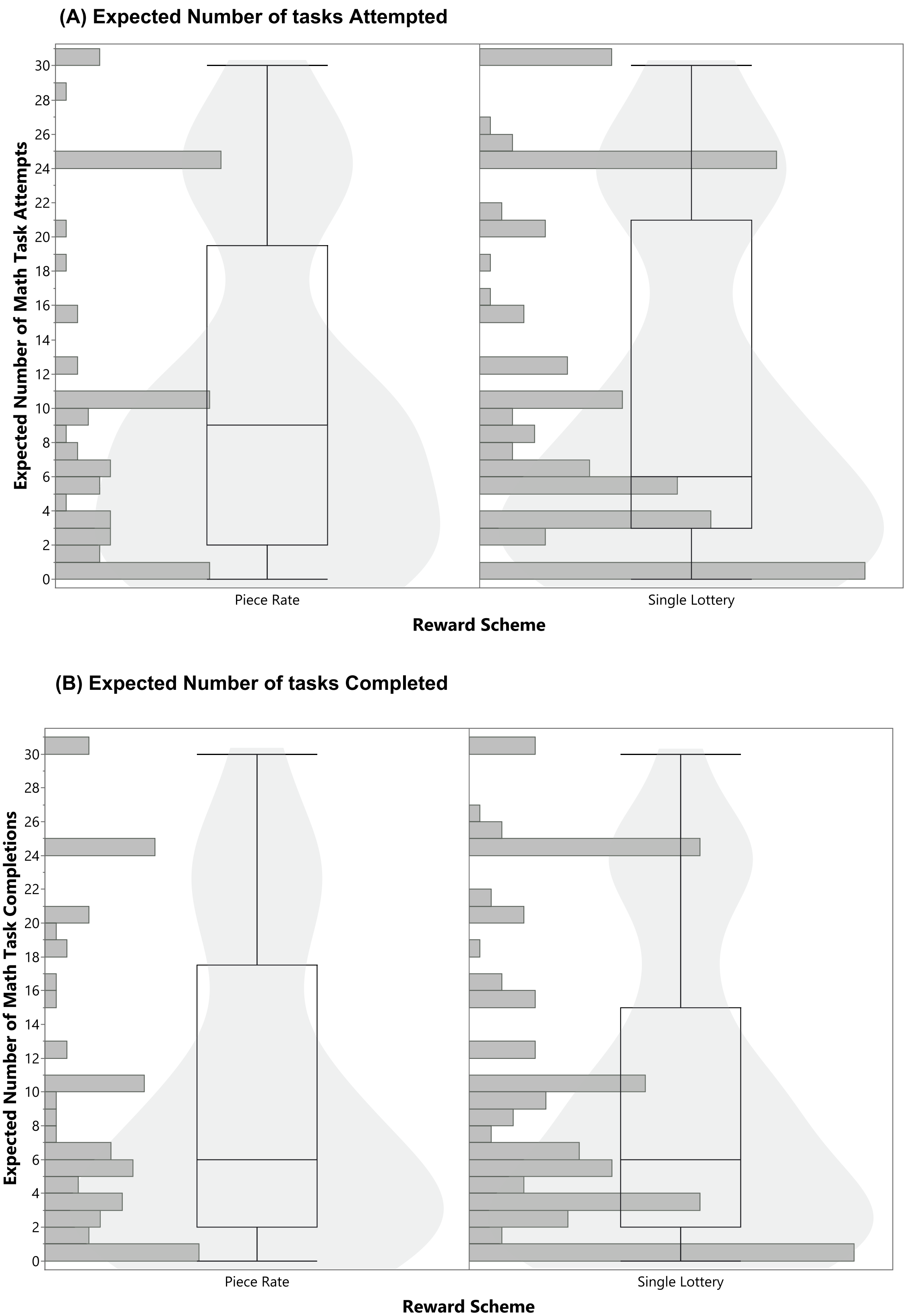

Fig. 4. Expected Motivation and Performance in Experiment 3. Note: Shared area represents violin plots.

based on the participant's expected number of task completions. We excluded individuals who indicated that their expected number of task completions would be less than 3 (corresponding to a 100% chance of $0 bonus) or more than 23 (corresponding to a 100% chance of $0.16 bonus). The predicted value of winning the lottery from the regressions is depicted by the bold line. Second, the optimism score for each participant was generated by taking their predicted value less the objective probability of winning the bonus based on their expected

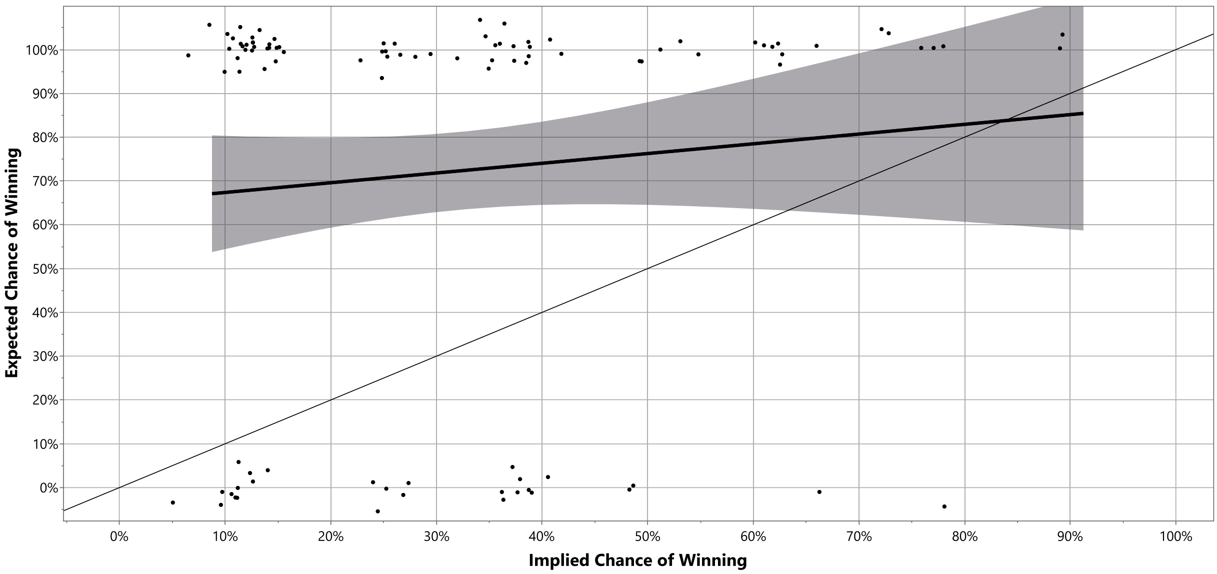

Fig. 5. Estimated Optimism Bias in Experiment 3. Notes: The black dots along the 0% and 100% value of the Y-axis (expected chance of winning) represents the number of people who expected to win no bonus and 100% bonus respectively. The percentage of people who expected to win full bonus when objective chance of winning was 12.5% was 71.85. For each subsequent 12.5% increase in objective probability the corresponding expected chance of winning was 68.6 (25), 65.2 (37.5), 66.7 (50), 88.9 (62.5), 83.3 (75), 100 (87.5) & 96.8 (100). The 45-degree line represents the case where there is no optimism bias. The bold line represents the estimated chance of winning based on responses of participants under the SL scheme (shaded area represents the confidence interval). The difference between the two lines is the measure of optimism bias.

number of task completions. To test whether there was an optimism bias, we conducted a *t*-test to examine whether the mean optimism score was different from zero. The mean optimism score was 0.40 and was statistically significant ($p < 0.001$), which means that participants were, on average, expecting to win the lottery with 40.2% higher chance than the objective probability.

To examine whether the optimism bias varied across expected performance, we ran an OLS regression where the dependent variable was the optimism bias, and the independent variable was expected performance. Our regression estimate (reported in Table 3) showed a negative and statistically significant relation between optimism bias and expected performance ($p < 0.001$): for every 1 unit increase in expected performance, optimism bias reduced by 3.2 percentage points.

Table 3
Summary of Expected Attempts, Expected Completions, and Optimism in Experiment 3.

	Expected Motivation		Expected Performance		Optimism
	(1)	(2)	(3)	(4)	(5)
SL	0.647	3.434	0.740	1.958	
	(0.531)	(3.344)	(0.490)	(3.087)	
Ability		1.516		1.332	−0.016
		(0.980)		(0.905)	(0.014)
SL*Ability		−1.240		−0.834	
		(1.168)		(1.078)	
Constant	8.814***	6.447*	7.287***	5.911*	0.445***
	(1.553)	(2.761)	(1.432)	(2.548)	(0.040)
Observations	263	263	263	263	101
R-squared	0.006	0.010	0.009	0.011	0.014

SL = Single lottery. Coefficient estimates from Probit regressions. Standard errors are in parenthesis. DV = Expected number of math task attempts in columns (1) and (2), expected number of math task completions in columns (3) and (4), and optimism in column (5). Regressions control assignment into SL or PR (columns 1–4). In addition, ability and its interaction with incentive schemes are controlled for in columns 2, 4, and 5. Ability was measured based on responses to the Berlin Numeracy Task. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

4.3. Discussion

The results of Experiment 3 support the hypothesis that part of the reason for the superiority of the single lottery reward scheme is optimism bias. Even though participants understood, based on their expected performance, the objective probability of winning the bonus, they systematically believed *they* were much more likely to win the lottery. To take the starkest example, one of the comprehension questions required participants to correctly indicate before progressing that "If you have correctly completed a total of 3 cognitive tasks, you will have a 12.5% chance of winning the $0.16 bonus payment". However, mere seconds later, 21 out of the 25 participants (84%) who expected to complete a total of 3 cognitive tasks indicated that they expected to win the bonus. The optimism bias was prevalent in our data, particularly for those expecting below average performance.

Although we believe that all probabilistic reward schemes initially benefit from the optimism bias, we argue that the single lottery reward scheme benefits the most. This is because the single lottery scheme is associated with just one lottery at the end of the reward period. As a result, workers can generate motivation from their optimism and, crucially, maintain that optimism-driven motivation for the entire period. By contrast, the multiple lottery scheme is associated with repeated lotteries. As a result, many workers will quickly face the reality of losing a lottery. For example, in Experiment 2, by design, 87.5% of participants in the multiple lottery probabilistic reward scheme group who completed 3 math tasks lost their first lottery. Research in the domain of experience- versus description-based risky choices suggests that, over time, the impact of experience overwhelms the impact of descriptions on choice (Weiss-Cohen et al., 2016). In our structure, the multiple lottery scheme is a type of experience-based choice because participants get repeated feedback from the lottery draws. Optimism is unlikely to thrive in the context of repeated losses.

In order to pursue this line of thinking, we ran a pre-registered study with 313 participants recruited from AMT comparing the SL scheme to the ML scheme as well as a version of the ML scheme where the earned lotteries were stockpiled and played back-to-back near the end of the study. By using this design those under this revised ML scheme – what we call the ML "end", or ML-E, scheme – were free to maintain their optimism for the entire period, similar to those under the SL scheme. To increase the generalizability of our findings, we used a new letter counting cognitive task (see Rosaz & Villeval, 2012). Replicating the findings of Experiments 1 and 2, motivation was significantly higher for those under the SL scheme than those under the ML scheme. Motivation for those under the ML-E scheme was somewhere in the middle of the SL and ML schemes. Details can be found in Appendix E: Supplementary Study 3. These findings lend support to the idea that one of the attributes driving the superiority of the SL scheme is that the risk uncertainty resolution occurs at the end of the period. As a result, workers can maintain their optimism and not be discouraged by losing the lottery.

5. Experiment 4

The purpose of pre-registered Experiment 4 was to shed light on another mechanism driving the effectiveness of the single lottery probabilistic reward scheme. According to our theorizing, the single lottery scheme uniquely benefits from goal progress and the goal gradient effect since it is possible to obtain all lottery tickets and guarantee reaching the goal. This potential to eliminate the uncertainty – which is absent from the multiple lottery probabilistic reward scheme – is critical because, without it, the motivational link between actions (completing work; in this case, solving math puzzles) and outcomes (monetary reward) is weakened. To test this aspect of the theorized mechanism, we compared two single lottery schemes in which superior performance either could or could not guarantee the lottery win. In addition, to address H4, we asked questions pertaining the perceptions of the schemes.

5.1. Methods

The experiment was carried out online. The 340 participants (39% male; $M_{\text{age}} = 19.4$ years) were recruited from a university student subject pool. Each participant received course credit for completing the study as well as a bonus payment continent on their decisions and probabilistic outcomes during the study.

The study used the same real effort task as in Experiment 2 followed by a questionnaire. Participants were randomly allocated to one of three bonus reward schemes: (1) Piece rate (PR) where, for every 3 math tasks completed, a bonus of $0.04 was earned, with a maximum bonus of $0.32; (2) Single Lottery (SL) where, for every 3 math tasks completed, an additional 12.5% chance of winning a lottery worth $0.32 was accrued; and (3) Single Lottery - Uncertainty (SL + U) where, for every 3 math tasks completed, an additional 5% chance of winning a lottery worth $0.80 was accrued. Note that the maximum possible chance of winning for those in the SL group versus SL + U group was 100% versus 40%, respectively. Note also that the lottery bonus for those in the SL group versus SL + U group was $0.32 versus $0.80, respectively. As a result, the expected value between schemes was equal. Given the different sample, we doubled the incentives compared to Experiment 2 based on pilot data and a goal to avoid ceiling effects. Following our pre-registration plan, we oversampled the PR and SL groups because of our intention to examine their interaction with ability.

The follow-up questionnaire was the same as that used in Experiment 2 with the addition of several fairness questions similar to those used in Experiment 1. In particular, we measured the participant's perceived difficulty to earn the bonus (1 = "Extremely easy"; 7 = "Extremely difficult"), perceived effectiveness their bonus scheme if the participant was a manager tasked with motivating employees (1 = "Ineffective"; 7 = "Effective"), perceived distributive justice ("How fair was the bonus you received considering the amount of effort that you put into your work?"; 1 = "Extremely unfair"; 7 = "Extremely fair"), and perceived procedural justice ("How fair was the procedure used to determine your bonus?"; 1 = "Extremely unfair"; 7 = "Extremely fair"). On the next page, since the participants were familiar only with the scheme they had been allocated to, we explained the differences between the PR and SL schemes. We then asked which scheme the participant thought would more often lead other participants in the study to choose the math task. Finally, we asked which scheme the participant would use if they were a manager tasked with motivating both low-ability and high-ability workers. The average time to complete the experiment was approximately 43 minutes.

5.2. Results

With respect to motivation, averaging across groups and rounds, participants chose to attempt math task 77.3% of the time. However, as shown in Fig. 6, motivation varied across schemes. To analyze these data, as per our pre-registered analysis plan, we used a random-effects panel regression (probit link function, random intercept) to account for multiple observations per participant. The dependent variable was the choice to attempt the math task (0 = *No* and 1 = *Yes*) and the independent variables were the different incentive schemes. In Table 2, we present the estimated mean differences across groups. Supporting H1i, we found that those in the SL scheme were significantly more likely to choose the math task than in the PR scheme. More important for the purposes of this experiment, the attempt of math task was higher for those incentivized under the SL scheme than those in the SL + U scheme.

With respect to performance, averaging across groups and rounds, participants completed 12.75 out of 30 math tasks, which corresponds approximately to a 42.5% completion rate. To analyze these data, we used a regression framework where task completion was the dependent variable (0 = *Not completed* and 1 = *Completed*) and the independent variables were the different incentive schemes. Table 2 presents the mean differences across groups. Supporting H1ii, the completion rate under the SL scheme was significantly higher than that under the PR scheme. In addition, the completion rate under the SL scheme was significantly higher than under the SL + U scheme. We discuss the effects of ability in the *Heterogenous Effects Across Ability* section and the perceptions of fairness in the *Fairness and Effectiveness Perceptions* section.

5.3. Discussion

The results of Experiment 4 support the hypothesis that part of the reason for the superiority of the single lottery reward scheme is its potential to eliminate uncertainty. Research suggests that certain outcomes are perceived as categorically different and weighted more heavily than uncertain outcomes (Kahneman & Tversky, 1979). By completing 24 math tasks, those in the single lottery reward scheme were guaranteed a reward. Therefore, 24 was a strong reference point. For these participants, each completed math task produced increasingly more utility as this strong reference point was approached. In contrast, by completing 24 math tasks, those in the single lottery-uncertain reward scheme were guaranteed only a 40% chance of a reward. For them, 24 was a relatively weaker reference point.

6. Heterogeneous effects across ability

In this section, we investigated whether the ability of participants heterogeneously affected their motivation and performance across schemes in Experiments 1, 2, and 4. In Experiment 1, ability was measured by performance (number of sliders completed) in the Baseline round. In Experiments 2 and 4, ability was measured by performance in the Berlin Numeracy task. In Experiment 2, 32.1% of our AMT sample scored 1, 34.9% scored 2, 17.1% scored 3, and 15.9% scored 4. In Experiment 4, 29.0% of our student sample scored 1, 37.3% scored 2, 19.8% scored 3, and 13.9% scored 4.

To analyze the heterogeneous impact of ability on motivation and performance across schemes, we utilized similar regression frameworks described in the experiment report sections to generate mean group differences. For Experiment 1, the dependent variable was the number of slider tasks completed in the first scheme encountered and the independent variables were the schemes and their interaction with the

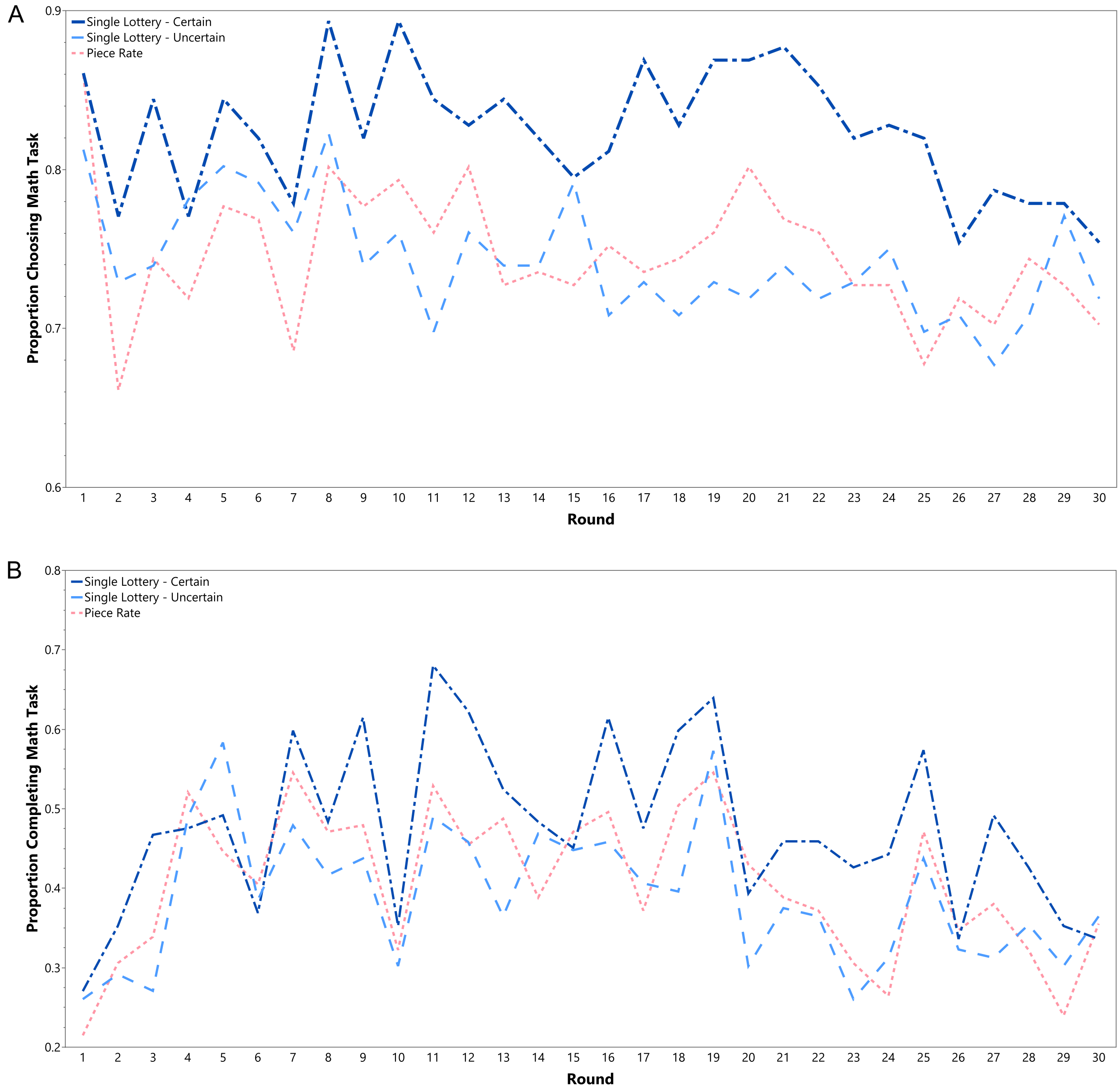

Fig. 6. (A) Proportion Choosing to Complete the Math Task and (B) Proportion of Completing the Math Task Split by Round and Incentive Scheme in Experiment 4.

ability. For Experiments 2 and 4, the dependent variables were either choice to attempt the math task (0 = *No* and 1 = *Yes*) or correct completion of the math task (0 = *No* and 1 = *Yes*). The independent variables were reward scheme and its interaction with ability. In each regression, we limited the sample to participants in the PR scheme and one of the probabilistic reward schemes.

The results are reported in Table 4. In Experiment 1, we observed that ability had a positive and significant effect on performance in the ML and PR schemes, but not in the SL scheme. However, the interactions were not significant indicating that neither the SL nor ML scheme produced higher performance for those with relatively lower (vs. higher) ability. In Experiment 2, we again found that ability had a significant

Table 4
Heterogenous Effect of Ability on Motivation and Performance in Experiments 1, 2, and 4.

	Experiment 1				Experiment 2			Experiment 4
	Performance		Motivation		Performance		Motivation	Performance
Effect of ability in SL scheme	0.452		−0.010		0.048		0.033	0.066***
Effect of ability in ML scheme		1.170***		0.101**		0.092***		
Effect of ability in PR scheme	0.831***	0.831***	0.070*	0.080*	0.129***	0.126***	0.036	0.033
Additional impact of ability in the SL schemes	−0.380		−0.080		−0.081		−0.003	
Additional impact of ability in the ML schemes		0.343		0.021		−0.034		0.032

PR = Piece Rate, SL = Single lottery, ML = Multiple lottery. Average marginal effects of ability for each subsample based on random effects Probit regression is presented. In all regression specifications, sample from two schemes (either ML or SL vs. PR) was utilized. The additional impact presents the marginal effect of coefficient of the interaction term in the regressions. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

positive effect on the choice to attempt as well as complete math tasks in the PR and ML schemes, but not the SL scheme. However, we again found no significant interaction between ability and the probabilistic scheme. Finally, in Experiment 4, we found no significant effect on the choice to attempt as well as complete math tasks in the PR scheme. By contrast, we did find a significant positive effect of ability on correct completions (but not choice of attempt) in the SL scheme. In sum, these results show that the effect of ability is less pronounced in the SL scheme relative to the PR scheme. However, the additional impact of ability was never statistically different across the PR and the probabilistic schemes.

Given the mixed findings for worker ability as a moderator of the main effect, we ran a new pre-registered study with 323 participants. The impetus of this study was recognition that the measure of ability used in Experiments 2 and 4 (i.e., performance on the Berlin Numeracy test) may only be tangentially related to the focal task of these experiments (i.e., quickly adding numbers together). Therefore, in this new study, we measured ability using a task nearly identical to the focal task, and did so more than 1 week before the focal task. We reasoned that this would be a stronger test of the worker ability boundary effect hypothesis. Once again, the evidence was somewhat inconclusive. In sum, our pre-registered analysis plan yielded no significant differences; however, we did see a significant interaction between reward scheme and worker ability in exploratory analyses. These exploratory analyses involved focusing on just the second half of the rounds (presumably the period during which lower ability workers came to more accurately perceive their objectively low ability level) while other exploratory analyses involved treating ability as a binary rather than continuous variable. See Appendix D: Supplementary Study 2 for details.

Finally, to get a clearer picture of the impact of ability, we conducted a single paper *meta*-analysis of Experiments 1, 2, 4 and Supplementary Study 2 following the protocol set by McShane and Böckenholt (2017). For the motivation analysis, we used data from Experiments 2, 4, and Supplementary Study 2. For the performance analysis, we used data from Experiments 1, 2, 4, and Supplementary Study 2. The omission of Experiment 1 from the motivation analysis was due to the nature of the outcome of the task used, which cannot be broken down into motivation and performance components. In our judgement, the dependent variable observed in Experiment 1 (the correct completion slider-tasks) aligned better with our definition of performance in the subsequent studies. In all our analysis, we restricted ourselves to the PR and the SL schemes.

For the motivation analysis, the task used across the studies was the same (i.e., 30 rounds of the math task) and generated a total of 18,960 observations. The dependent variable was the choice to complete the math task (0 vs. 1) and the two independent variables were the bonus scheme (PR vs. SL), ability (high vs. low), and the interaction term. For the performance analysis (n = 19,019), the correct completion of the math task was used as the dependent variable for the math-task based studies. In addition, for Experiment 1, the number of completed slider task was the measure of performance.

Across studies, ability was measured differently. For Experiments 2 and 4, a participant was classified as high ability if they scored greater than 2 (out of 4) in the Berlin Numeracy Task. Based on this classification, ~34% of the participants in each of Experiment 2 and 4, were classified as having high ability. Supplementary Study 2 was a two-stage study. In the first stage, participants were asked to correctly complete as many math tasks as possible in 10 minutes. The performance in the first stage was used as a continuous measure of ability. A similar strategy was utilized in Experiment 1. For our purposes, we conducted median splits of performance across participants in these experiments and classified anyone who performed better than the median as high ability (~48% and 56% of the sample for Supplementary Study 2 and Experiment 1, respectively).

Using the data, we generated five contrasts: (i) effect of high-ability under the PR scheme, (ii) effect of high-ability under the SL scheme, (iii) effect of SL (vs. PR) treatment on high-ability participants (iv) effect of SL (vs. PR) treatment on low-ability participants, and (v) the additional effect of high-ability under the SL scheme relative to the PR scheme (i.e., interaction effect). Table 5 reports the results. First, high (vs. low) ability participants showed 16.4%-point more motivation and 50.8%-point higher performance under the PR scheme. By contrast, under the SL scheme, high (vs. low) ability participants were not statistically more motivated nor demonstrated higher performance. Second, for low ability workers, the SL (vs. PR) scheme was more motivating and generated higher performance. By contrast, for high-ability workers, the bonus scheme did not impact motivation. Note that both these findings are consistent with our H3 hypothesis regarding worker ability. Finally, the additional impact of the SL (vs. PR) scheme between high-ability and low-ability workers (i.e., the interaction effect) was negative. The point-estimate of −0.114 and −0.193, suggest the impact of ability on motivation (performance) was 11.4 (19.3) %-point higher under the PR scheme than under the SL scheme. However, these estimates were imprecise and did not reach statistical significance. Taken together, the conclusion we can make is that the SL scheme is better at motivating and generating higher performances relative to the PR scheme for low-ability workers.

7. Fairness and effectiveness perceptions

In two of the experiments, we asked participants about perceived fairness and effectiveness of the various schemes that they participated in. In Experiment 1, participants were asked to indicate how fair and how effective the four different schemes they experienced were. In Experiment 4, participants were asked to indicate the fairness of the bonus they received relative to their effort (distributional justice), the process by which the bonus was determined (procedural justice), and the effectiveness the reward scheme. Participants in Experiment 4 were also provided with a description of the scheme they had not directly participated in and asked (i) to indicate which scheme was better at motivating math task attempts and (ii) to choose between the SL and PR scheme to motivate low-skilled and high-skilled workers.

7.1. Fairness perceptions

In terms of perceived fairness, as shown in Table 6, participants in Experiment 1 judged the PR scheme to be fairer than the SL and ML schemes. The LS scheme was also judged to be fairer than the ML scheme. Participants in Experiment 4 judged the PR scheme to have higher distributional and procedural fairness than the SL scheme. In sum, and in support of H4, we found clear evidence for the probabilistic schemes to be perceived as less fair than the non-probabilistic schemes. Notably, of the two probabilistic schemes, the SL scheme was judged as

Table 5
Single Paper Meta-Analysis Results of Incentive Scheme and Ability on Motivation and Performance.

	Motivation		Performance	
	Diff.	*p*-value	Diff.	*p*-value
	(1)	(2)	(3)	(4)
Effect of high-ability (vs. low) under PR scheme	0.164	0.008	0.508	0.006
Effect of high-ability (vs. low) under SL scheme	0.050	0.303	0.315	0.085
Effect of SL (vs. PR) scheme on high-ability participants	0.031	0.522	0.186	0.307
Effect of SL (vs. PR) scheme on low-ability participants	0.145	0.003	0.379	0.039
Additional effect of high-ability under the SL scheme relative to the PR scheme (i.e., interaction effect)	−0.114	0.098	−0.193	0.456

PR = Piece Rate; SL = Single lottery. Difference estimates (columns 1 and 3) and *p*-values (columns 2 and 4) are generated using the procedure outlined in McShane & Böckenholt (2017).

Table 6
Summary of Post-Experiment Questionnaire on Fairness and Effectiveness in Experiments 1 and 4.

	Fairness			Effectiveness	
	Experiment 1	Experiment 4		Experiment 1	Experiment 4
		Distributive	Procedural		
Panel A: Mean Values					
SL	4.59	3.14	3.95	4.85	4.10
ML	3.67			3.62	
PR	5.22	3.78	4.45	5.08	4.38
LS	4.87			5.71	
Panel B: Difference Across Schemes					
SL vs. PR	0.63**	0.64**	0.50*	0.23	0.29
SL vs. LS	0.28			0.86***	
SL vs. ML	0.92***			1.22***	
ML vs. PR	1.55***			1.46***	
ML vs. LS	1.20***			2.09***	

LS = Lump sum, PR = Piece Rate, SL = Single lottery, ML = Multiple lottery. Panel A presents mean values of fairness (out of 7) and effectiveness (out of 7). Panel B presents differences between incentive schemes where *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ based on *t*-tests.

fairer.

To further investigate fairness perceptions, we asked 267 new participants to evaluate procedural and distributional justice after seeing the performance and bonus outcomes of a (simulated) company using the SL scheme after 1 month and after 1 year. We reasoned that participants in Experiments 1 and 4 were overly focused on the outcomes of the SL scheme after one period (and, hence, one lottery) and might change their opinion over time when multiple lotteries had been conducted. This is because under the SL scheme the correlation between performance and bonus increases with each additional lottery. Consistent with our pre-registered hypothesis, fairness perceptions of the SL scheme were significantly better when reflecting on the performance and bonuses of workers in the company after 1 year (i.e., 12 lotteries played) than 1 month (i.e., 1 lottery played). See Appendix C: Supplementary Study 1 for details.

7.2. Effectiveness perceptions

In terms of perceived effectiveness, as shown in Table 6, participants in Experiment 1 evaluated the LS scheme as more effective than the SL and ML schemes. They also evaluated the PR scheme to be more effective than the ML scheme. Participants in both Experiments 1 and 4 found the PR and SL schemes to be equally effective. In sum, we found mixed evidence for the perceived effectiveness of the various probabilistic and non-probabilistic schemes. Notably, of the two probabilistic schemes, the SL scheme was evaluated as more effective.

In terms of inclination towards using the different schemes, as shown in Table 7, participants believed the PR (vs. SL) scheme was better at motivating others in the study. When asked to assess the PR and the SL scheme from a manager's perspective, participants leaned toward the PR scheme for low-ability workers and were indifferent between schemes for high-ability workers. In sum, we found evidence suggesting that participants believed the PR scheme would be preferred, particularly for low-ability workers.

Table 7
Participant's Beliefs about Motivation Produced by SL and PR Schemes in Experiment 4.

	N	Overall Motivation	Choice of Scheme as Manager for		
			Low-ability Workers	High-ability Workers	*p*-value[a]
Full Sample	339	42.5***	39.6***	48.7	0.05
PR Group	120	42.5	48.3	44.2	0.60
SL Group	122	45.1	34.7***	50.8	0.04
SL + U Group	97	39.2*	35.5**	51.6	0.05
p-value[b]		0.68	0.05	0.47	

PR = Piece Rate, SL = Single lottery, SL + U = Single lottery + uncertainty. Each number represents the proportion of participants in that sample that believed the SL scheme was better at motivating others than the PR scheme. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ based on two-sided binomial test examining the hypothesis that the proportion of participants choosing SL scheme is 0.5. *p*-value. [a] based on a two-sided *t*-test to investigate differences in choices of participants for low-ability and high-ability workers. [b] based on a two-sided χ^2 test to investigate differences in choices of participants from the three different subsamples.

8. General Discussion

The main research question addressed in this work was whether a probabilistic reward scheme could increase worker motivation and performance compared to a more traditional, non-probabilistic reward scheme. Our research suggests that the answer is *yes* – but it depends critically on the type of probabilistic scheme. Across three different tasks varying in cognitive complexity, participants displayed higher motivation and produced more output when they were incentivized with a unique type of probabilistic reward scheme – where the association between performance and reward was uncertain – than when rewards were allocated with certainty after a pre-specified level of performance. These observations complement findings from finance (Filiz-Ozbay et al., 2015), marketing (Mazar et al., 2017), pro-sociality (Diamond & Loewy, 1991), and health (Volpp, John, et al., 2008) domains indicating that probabilistic rewards, when designed effectively, can predictably influence behavior.

Importantly, it turns out that not all probabilistic rewards schemes are equal. We contrasted two types of probabilistic rewards scheme: one in which additional performance increased the number of lotteries the worker participated in (the "multiple lottery" scheme) with one in which additional performance increased the probability of winning a one-off lottery at the end of the work period (the "single lottery" scheme). Our research revealed that the single lottery probabilistic scheme produced the most motivation and the best performance. This finding is consistent with past research that found no benefit of the multiple lottery scheme compared to traditional schemes (Haisley, 2008).

8.1. Theoretical implications

Overall, these results challenge the assumption of theories which assert that people are averse to uncertainty (Gneezy et al., 2006). In fact, there are several reasons why a probabilistic reward scheme is motivating. We found support for the idea that many people are optimistic about their chances of winning (Helweg-Larsen & Shepperd, 2001).

Such optimism is reinforced by people's tendency to overweight small, described probabilities (Kahneman & Tversky, 1979). Indeed, we found that the greatest degree of optimism was displayed by those with relatively lower chances of winning. This optimism bias is likely to help overcome the "starting problem", which refers to the difficulty in getting started on a difficult goal (Heath et al., 1999). Optimism is such a powerful tendency that we speculate that even if a worker does not win the probabilistic reward, they can easily maintain the belief that they will win the next period. Such a belief is likely to be reinforced by the gambler's fallacy (Clotfelter & Cook, 1993): the observation of a colleague winning the bonus may all but guarantee in the worker's mind that they will win the next lottery because it is "their turn".

Although we focused on a reward program implemented during a single work period (i.e., in which the single lottery was played just once), the actual implementation of this scheme would repeat across periods (e.g., each new month). Over repeated periods, a probabilistic reward scheme is likely to benefit from learning according to the principles of operant conditioning (Ferster & Skinner, 1957). Of particular interest is the schedule of reinforcement, which refers to the rule by which instances of behavior are reinforced or punished. Similar to a probabilistic reward scheme, a variable-ratio schedule is one in which a reward is given after an average number of actions have been made; however, the specific action upon which the reward is provided is uncertain. Classic work has shown that variable-ratio, compared to other schedules, produce higher rates of action in both animals and humans (Pritchard et al., 1976; Zuriff, 1970).

The difference in observed behavior between the two probabilistic reward schemes is perhaps the most theoretically insightful contribution. In Table 8, we outline the different features of the schemes we examined. The single lottery scheme, which was superior to the multiple lottery scheme, possesses four unique features that we believe explain its superiority.

1. Reference point: The single lottery scheme creates a strong reference point in terms of the performance required to collect all the lottery tickets, which is within reach of at least the high-ability workers (Heath et al., 1999). As a result, this scheme benefits from motivation derived from goal progress (e.g., Nunes & Dreze, 2006) and capitalizes on the goal-gradient effect: the additional motivation people are willing to expend to close the gap between a goal and just missing the goal (Kivetz et al., 2006). By contrast, the multiple lottery scheme creates multiple weak reference points related to the performance required to trigger the next lottery.
2. Risk reduction: The single lottery scheme creates a goal-pursuit framework wherein additional work reduces the uncertainty. By contrast, the multiple lottery scheme simply provides another attempt at the same lottery. Indeed, under the multiple lottery scheme, increased performance is associated increasing variance in potential rewards.
3. Risk elimination: The single lottery scheme that we implemented allowed participants to achieve certainty if the goal was attained. Much research points to the additional value that people accrue to certainty compared to even slightly-less-than-certain outcomes (Kahneman & Tversky, 1979). By contrast, even high-ability workers can end up with nothing under the multiple lottery scheme. By permitting the possibility of a guaranteed reward under the single lottery scheme (by accumulating all lottery tickets), this scheme combines the best features of probabilistic and non-probabilistic rewards schemes: The probabilistic element allows motivation to be derived from optimism and excitement and the non-probabilistic element allows motivation to be derived from the certain and very fair link between performance and reward. The possibility of certainty under the single lottery scheme may even decrease worker stress, which can undermine motivation and performance (Cadsby et al., 2007).
4. Risk resolution timing: The single lottery scheme resolves the risk uncertainty at the end of the period whereas the multiple lottery scheme resolves the risk uncertainty during the period. As a result, only those under the single lottery scheme can benefit from optimism for the entire duration of the task. By contrast, those under the multiple lottery are likely to have their motivation undermined by the experience of losing a lottery.

We acknowledge that our instantiation of the multiple lottery scheme resulted in frequent experiences of losing the lottery. This was done in order to equate the lottery win amount under the single lottery and multiple lottery schemes. However, experiencing frequent losses is not a core feature of the ML scheme. An alternative ML scheme could make winning the lottery the more common experience. However, in order to equate expected values, the lottery amount under the single lottery scheme would need to increase. We leave this titration act between win percentage and win amount to future work.

We had hypothesized that probabilistic schemes would have most beneficial impact on low-ability workers who are likely to produce less than the median across all workers. This is because traditional non-probabilistic reward schemes typically benefit high-ability workers and can undermine the motivation of low-ability workers because of their inability to realistically obtain the bonus. Consistent with this analysis, a recent study examining the effect of changing an organization's incentive schemed *away* from performance-related pay towards higher fixed pay found a drop in productivity for high-ability workers (Bun & Huberts, 2018). By contrast, probabilistic schemes – in particular, the single lottery scheme – provide low-ability workers with a realistic opportunity to obtain a reward just as large as high-ability workers, albeit with a smaller likelihood. Consistent with this

Table 8
Attributes of the Different Incentive Schemes.

Attribute	Description	Scheme					
		Piece Rate	*Lump sum*	*Multiple Lottery*	*Multiple Lottery (End)*	*Single Lottery (Uncertainty)*	*Single Lottery*
Ascendance	The expected value of the reward increases per unit of performance	✓	×	✓	✓	✓	✓
Reference Point	There is a strong reference point that can naturally become a performance goal	×	✓	×	×	✓	✓
Risk	It is possible to receive a reward less than the expected value	×	×	✓	✓	✓	✓
Risk Reduction	It is possible to reduce the risk uncertainty through higher performance	*N/A*	*N/A*	×	×	✓	✓
Risk Elimination	It is possible to entirely eliminate the risk uncertainty through high performance	*N/A*	*N/A*	×	×	×	✓
Risk Resolution	The risk uncertainty is resolved at the end of the period	*N/A*	*N/A*	×	✓	✓	✓

theorizing, we found that those expecting relatively low performance displayed the strongest optimism bias under the single lottery scheme. Averaging across all our studies, it was the low ability workers who were most affected by the single lottery scheme.

8.2. Practical implications

Our results have important implications for how managers design incentive reward schemes. One implication is that managers should be cautious when relying on their own intuition regarding the best incentive scheme. This is because we observed that participants most strongly preferred the lump sum scheme and believed it to be the most effective. However, this belief was in spite of the fact that the lump sum system was objectively the *least* effective reward system of the four trialed in Experiment 1. There was also some evidence that the piece rate scheme was perceived to be at least as, if not more, effective than the single lottery scheme despite our results suggesting the reverse to be true.

We believe it is important to emphasize that it may be quite foolish for a manager to change their existing non-probabilistic bonus scheme to a probabilistic one. The backlash is likely to be swift and powerful, as was experienced by United Airlines in 2018 after they announced a plan to replace automatic modest bonuses for all with a lottery to win large prizes for just a few (Cancialosi, 2018). Within a few days of the announcement, United Airlines scrapped the plans.

This backlash can be better understood considering the fairness perceptions that we observed. On average, probabilistic schemes were considered less fair than non-probabilistic schemes. Compared to a piece rate scheme, the single lottery scheme was perceived to be a less fair process for distributing rewards and produced less fair bonus allocations (Cappelen et al., 2007; Deutsch, 1975; Lind & Tyler, 1988). Importantly, we found that the perceived unfairness of probabilistic reward schemes dissipated as people recognized the strong positive correlation between performance and bonus in the long run. Future research could investigate additional ways to ameliorate fairness concerns. One idea is to increase transparency. Research has revealed that providing workers with information about their relative performance increases their motivation (Cadsby et al., 2019; Hannan et al., 2008), and this effect is stronger when the relative performance information is communicated publicly versus privately (Hannan et al., 2012; Tafkov, 2012).

Our managerial recommendation is not to replace or restructure an existing reward scheme with a probabilistic one because such a change would likely be perceived – as exemplified by United Airlines employees – as a loss. After all, an employees' paycheck would suddenly become uncertain. Rather, a probabilistic reward scheme may be a strong candidate for any manager planning to introduce a contingent reward scheme for the first time. Such a scheme should not threaten current financial incentives but, instead, be "layered" on top as bonus compensation. According to the theoretical analysis by Wiseman and Gomez-Mejia (1998), such a layering approach is unlikely to be perceived as a threat to perceived wealth.

An important consideration for a manager is the type of work that is being incentivized, which is likely to interact with the effectiveness of the incentive structure (Bonner et al., 2000). In general, pay-for-performance and, thus, any associated probabilistic reward scheme, is most appropriate when employee output is independent, easy to divide into sub-goals, and both cheap and easy to measure objectively (Brown, 1990). Examples of such work contexts include arts (e.g., photographer), clerical (e.g., typist), computing (e.g., technician), design (e.g., tattooist), logistics (e.g., courier), manufacturing (e.g., sewing machinist), retail (e.g., car salesperson), and transport (e.g., taxi driver). As work becomes more cognitively complex and reliant on intrinsic motivation, a probabilistic reward scheme is less likely to improve performance and so we would not recommend it (Cerasoli et al., 2014). Examples of such work contexts include business management (e.g., management consultant), construction (e.g., construction manager), education (e.g., teacher), engineering (e.g., mechanical engineer), financial services (e.g., economist), healthcare (e.g., nurse), hospitality (e.g., hotel manager), legal services (e.g., lawyer), performing arts (e.g. actor), and science (e.g., chemist).

8.3. Limitations and future directions

Our research was limited to contexts in which the individual was evaluated only in light of a pre-specified performance target and participated in their own personal lottery for a relatively small individual reward in a controlled lab study. Although we relied on a convenience sample, participants recruited from Amazon's Mechanical Turk are on the platform to earn money in exchange for carrying out tasks, so the labor context we studied is quit apt. It would be worthwhile to examine how well probabilistic rewards motivate in the context of a team-based scheme – where probabilistic rewards were tied to a team, rather than individual, performance – and what conditions were optimal (e.g., size and gender composition of the team; equitably vs. equally distributed rewards; Garbers & Konradt, 2014). It would be also interesting to examine how probabilistic rewards motivate in the context of a tournament-based scheme – where probabilistic rewards were tied to relative, rather than absolute, performance – and what conditions were optimal (e.g., tournament size, prize structure, handicapping; Connelly et al., 2014). Given that extrinsic rewards can undermine intrinsic motivation, another interesting future direction is to investigate to what extent probabilistic rewards motivate when part or all of the lottery rewards are donated to a cause the individual cares about. Finally, it would be valuable to discover what organizational design elements complement the usage of a probabilistic reward scheme (e.g., decentralization of decision making; Hong et al., 2019) as well as explore the consequences of probabilistic rewards on broader values and behavior (Hur et al., 2021).

In this paper, we focused on the individual difference variable of ability. Another potentially important individual difference variable is financial resources. On the one hand, those with considerable financial resources may be more open to a probabilistic reward scheme because the outcome is financially inconsequential, which is consistent with the observation of higher risk taking in high-income, high-wealth households (Fang et al., 2021). On the other hand, those with few financial resources may be more open to a probabilistic reward scheme because of a strong hope or even need to win the lottery, which is consistent with problem gambling behavior (Castrén et al., 2018).

Future research could also examine how tight the link between performance and probabilistic reward needs to be. In our study, there was a relatively strong positive correlation. However, some recent high-profile examples have implemented much weaker correlations. For example, in 2022, the fast-food chain Raising Cane's bought US $200,000 worth of lottery tickets on behalf of its 50,000 employees with the promise that a lottery win would be shared equally between all employees (Paúl, 2022). Understanding to what extent this kind of probabilistic reward, which is not linked to an individual worker's performance, affects motivation in the short and long-term and for what reasons is a worthwhile future research question.

8.4. Conclusion

In today's tight job market, it is becoming increasingly challenging to attract, retain, and motivate employees. The current research indicates that the introduction of a probabilistic reward scheme – one in which each worker accumulates lottery tickets in their own personal lottery to win a moderate bonus – is perhaps one way to help address the problem. For managers looking to motivate and retain their employees, a single lottery probabilistic reward scheme may be worth the gamble.

CRediT authorship contribution statement

Adrian R. Camilleri: Conceptualization, Methodology, Software,

Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Katarina Dankova:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Jose M. Ortiz:** Conceptualization, Methodology, Software, Writing – review & editing. **Ananta Neelim:** Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Stimuli, data, and code are on OSF (https://osf.io/2ug97/?view_only=168452a61d9d430ea04fe547a7a1a4c3).

Acknowledgement

This research was supported by internal grants from the University of Technology Sydney, University of Tasmania, and RMIT University.

References

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience, 21*(8), 2793–2798.
Blinder, A. S. (2011). *Paying for productivity: A look at the evidence*. Brookings Institution Press.
Bonezzi, A., Brendl, C. M., & De Angelis, M. (2011). Stuck in the middle: The psychophysics of goal pursuit. *Psychological Science, 22*(5), 607–612.
Bonner, S. E., Hastie, R., Sprinkle, G. B., & Young, S. M. (2000). A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Management Accounting Research, 12*(1), 19–64.
Bonner, S. E., & Sprinkle, G. B. (2002). The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society, 27*(4–5), 303–345.
Brickman, P., Folger, R., Goode, E., & Schul, Y. (1981). Microjustice and macrojustice. In M. J. Lerner, & S. C. Lerner (Eds.), *The justice motive in social behavior: Adapting to times of scarcity and change*. Plenum Press.
Brown, C. (1990). Firms' choice of method of pay. *ILR Review, 43*(3), 165-S–82.
Bun, M. J., & Huberts, L. C. (2018). The impact of higher fixed pay and lower bonuses on productivity. *Journal of Labor Research, 39*(1), 1–21.
Cadsby, C. B., Song, F., Engle-Warnick, J., & Fang, T. (2019). Invoking social comparison to improve performance by ranking employees: The moderating effects of public ranking, rank pay, and individual risk attitude. *Journal of Economic Psychology, 72*, 64–79.
Cadsby, C. B., Song, F., & Tapon, F. (2007). Sorting and incentive effects of pay for performance: An experimental investigation. *Academy of Management Journal, 50*(2), 387–405.
Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty, 19*(1–3), 7–42.
Cancialosi, C. (2018). *Why United's Bonus Lottery Was Doomed From The Start*. Retrieved 22 November 2019 from https://www.forbes.com/sites/chriscancialosi/2018/03/12/why-uniteds-bonus-lottery-was-doomed-from-the-start/#6f1a78771692.
Cappelen, A. W., Hole, A. D., Sørensen, E.Ø., & Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review, 97*(3), 818–827.
Castrén, S., Kontto, J., Alho, H., & Salonen, A. H. (2018). The relationship between gambling expenditure, socio-demographics, health-related correlates and gambling behaviour—a cross-sectional population-based survey in Finland. *Addiction, 113*(1), 91–106.
Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin, 140*(4), 980.
Clotfelter, C. T., & Cook, P. J. (1993). The "gambler's fallacy" in lottery play. *Management Science, 39*(12), 1521–1525.
Clotfelter, C. T., Cook, P. J., Edell, J. A., & Moore, M. (1999). *State lotteries at the turn of the century: Report to the national gambling impact study commission* (Report to the National Gambling Impact Study Commission, Issue. https://pdfs.semanticscholar.org/5e7f/f036754660bd270c5c5eec3b5f801157d245.pdf.
Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes, 86*(2), 278–321.
Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy test. *Judgment and Decision Making, 7*(1), 25–47.
Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*(3), 386–400.
Condly, S. J., Clark, R. E., & Stolovitch, H. D. (2003). The effects of incentives on workplace performance: A meta-analytic review of research studies 1. *Performance Improvement Quarterly, 16*(3), 46–63.
Connelly, B. L., Tihanyi, L., Crook, T. R., & Gangloff, K. A. (2014). Tournament theory: Thirty years of contests and competitions. *Journal of Management, 40*(1), 16–47.
Cook, T., & Dixon, M. R. (2006). Performance feedback and probabilistic bonus contingencies among employees in a human service organization. *Journal of Organizational Behavior Management, 25*(3), 45–63.
Corazzini, L., Faravelli, M., & Stanca, L. (2010). A prize to give for: An experiment on public good funding mechanisms. *The Economic Journal, 120*(547), 944–967.
Corgnet, B., Gómez-Miñambres, J., & Hernán-Gonzalez, R. (2015). Goal setting and monetary incentives: When large stakes are not enough. *Management Science, 61*(12), 2926–2944.
DellaVigna, S., & Pope, D. (2018). What motivates effort? Evidence and expert forecasts. *The Review of Economic Studies, 85*(2), 1029–1069.
Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues, 31*(3), 137–149.
Diamond, W. D., & Loewy, B. Z. (1991). Effects of probabilistic rewards on recycling attitudes and behavior. *Journal of Applied Social Psychology, 21*(19), 1590–1607.
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.
Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review, 14*(1), 57–74.
Evans, K. M., Kienast, P., & Mitchell, T. R. (1988). The effects of lottery incentive programs on performance. *Journal of Organizational Behavior Management, 9*(2), 113–135.
Fang, M., Li, H., & Wang, Q. (2021). Risk tolerance and household wealth–Evidence from Chinese households. *Economic Modelling, 94*, 885–895.
Ferster, C. B., & Skinner, B. F. (1957). *Schedules of Reinforcement*. Appleton-Century-Crofts.
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.
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42. https://doi.org/10.1257/089533005775196732
Garbers, Y., & Konradt, U. (2014). The effect of financial incentives on performance: A quantitative review of individual and team-based financial incentives. *Journal of Occupational and Organizational Psychology, 87*(1), 102–137.
Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review, 102*(1), 469–503.
Gneezy, U., List, J. A., & Wu, G. (2006). The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *The Quarterly Journal of Economics, 121*(4), 1283–1309.
Goswami, I., & Urminsky, O. (2017). The dynamic effect of incentives on postreward task engagement. *Journal of Experimental Psychology: General, 146*(1), 1.
Guillén, M. F., & Tschoegl, A. E. (2002). Banking on gambling: Banks and lottery-linked deposit accounts. *Journal of Financial Services Research, 21*(3), 219–231.
Haisley, E. (2008). *The appeal of lotteries and their usage in incentive design*. Carnegie Mellon University].
Hannan, R. L., Krishnan, R., & Newman, A. H. (2008). The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *The Accounting Review, 83*(4), 893–913.
Hannan, R. L., McPhee, G. P., Newman, A. H., & Tafkov, I. D. (2012). The effect of relative performance information on performance and effort allocation in a multi-task environment. *The Accounting Review, 88*(2), 553–575.
Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology, 38*(1), 79–109.
Helweg-Larsen, M., & Shepperd, J. A. (2001). Do moderators of the optimistic bias affect personal or target risk estimates? A review of the literature. *Personality and Social Psychology Review, 5*(1), 74–95.
Hong, B., Kueng, L., & Yang, M.-J. (2019). Complementarity of performance pay and task allocation. *Management Science, 65*(11), 4951–5448.
Huang, S.-C., & Zhang, Y. (2011). Motivational consequences of perceived velocity in consumer goal pursuit. *Journal of Marketing Research, 48*(6), 1045–1056.
Hull, C. L. (1932). The Goal-Gradient Hypothesis and Maze Learning. *Psychological Review, 39*(1), 25–43.
Hur, J. D., Lee-Yoon, A., & Whillans, A. V. (2021). Are they useful? The effects of performance incentives on the prioritization of work versus personal ties. *Organizational Behavior and Human Decision Processes, 165*, 103–114.
Irwin, F. W. (1953). Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality, 21*(3), 329–335.
Isaac, J. E. (2001). Performance related pay: The importance of fairness. *Journal of Industrial Relations, 43*(2), 111–123.
Jenkins, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology, 83*(5), 777–787.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*(2), 263–292.
Kanfer, R. (1990). Motivation theory and industrial and organizational psychology. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (pp. 75–170). Consulting Psychologists Press.
Kanfer, R., & Chen, G. (2016). Motivation in organizational behavior: History, advances and prospects. *Organizational Behavior and Human Decision Processes, 136*, 6–19.

Kennedy, P. W. (1995). Performance pay, productivity and morale. *Economic Record, 71* (3), 240–247.
Kimmel, S. E., Troxel, A. B., French, B., Loewenstein, G., Doshi, J. A., Hecht, T. E., Laskin, M., Brensinger, C. M., Meussner, C., & Volpp, K. (2016). A randomized trial of lottery-based incentives and reminders to improve warfarin adherence: the Warfarin Incentives (WIN2) trial. *Pharmacoepidemiology and Drug Safety*, *25*(11), 1219-1227.
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.
Kirby, K. N., & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review, 3*(1), 100–104. https://doi.org/10.3758/Bf03210748
Kivetz, R. (2003). The effects of effort and intrinsic motivation on risky choice. *Marketing Science, 22*(4), 477–502.
Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research, 43*(1), 39–58.
Klein, H. J., Wesson, M. J., Hollenbeck, J. R., & Alge, B. J. (1999). Goal commitment and the goal-setting process: Conceptual clarification and empirical synthesis. *Journal of Applied Psychology, 84*(6), 885–896.
Landry, C. E., Lange, A., List, J. A., Price, M. K., & Rupp, N. G. (2006). Toward an understanding of the economics of charity: Evidence from a field experiment. *The Quarterly Journal of Economics, 121*(2), 747–782.
Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review, 90* (5), 1346–1361.
Lee, C.-Y., Morewedge, C. K., Hochman, G., & Ariely, D. (2019). Small probabilistic discounts stimulate spending: Pain of paying in price promotions. *Journal of the Association for Consumer Research, 4*(2), 160–171.
Lemieux, T., MacLeod, W. B., & Parent, D. (2009). Performance pay and wage inequality. *The Quarterly Journal of Economics, 124*(1), 1–49.
Lind, E. A., & Tyler, T. R. (1988). *The Social Psychology of Procedural Justice*. Springer Science & Business Media.
Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice-Hall Inc.
Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717.
Locke, E. A., Latham, G. P., & Erez, M. (1988). The determinants of goal commitment. *Academy of Management Review, 13*(1), 23–39.
Luyben, P. D., & Cummings, S. (1981). Motivating beverage container recycling on a college campus. *Journal of Environmental Systems, 11*(3), 235–245.
MacLeod, B. W., & Parent, D. (1999). Job characteristics and the form of compensation. *Research in Labor Economics, 18*, 177–242. https://doi.org/10.1108/S0147-9121(2012)0000035044
Maki, A., Burns, R. J., Ha, L., & Rothman, A. J. (2016). Paying people to protect the environment: A meta-analysis of financial incentive interventions to promote proenvironmental behaviors. *Journal of Environmental Psychology, 47*, 242–255. https://doi.org/10.1016/j.jenvp.2016.07.006
Mazar, N., Shampanier, K., & Ariely, D. (2017). When retailing and Las Vegas meet: Probabilistic free price promotions. *Management Science, 63*(1), 250–266.
McShane, B. B., & Böckenholt, U. (2017). Single-paper meta-analysis: Benefits for study summary, theory testing, and replicability. *Journal of Consumer Research, 43*(6), 1048–1063.
Mercer. (2018). *Global Compensation Planning Report*. https://www.imercer.com/products/global-comp-planning.aspx.
Meyer, H. H. (1975). The pay-for-performance dilemma. *Organizational Dynamics, 3*(3), 39–50.
Miller, M. V., Carlson, J., & Sigurdsson, S. O. (2014). Improving treatment integrity in a human service setting using lottery-based incentives. *Journal of Organizational Behavior Management, 34*(1), 29–38.
Mineka, S., & Hendersen, R. W. (1985). Controllability and predictability in acquired motivation. *Annual Review of Psychology*.
Moon, A., & Nelson, L. D. (2019). The uncertain value of uncertainty: When consumers are unwilling to pay for what they like. *Management Science*. https://doi.org/10.1287/mnsc.2019.3426
Nunes, J. C., & Dreze, X. (2006). The endowed progress effect: How artificial advancement increases effort. *Journal of Consumer Research, 32*(4), 504–512.
O'hora, D., & Maglieri, K. A. (2006). Goal statements and goal-directed behavior: A relational frame account of goal setting in organizations. *Journal of Organizational Behavior Management, 26*(1–2), 131–170.
Oberholzer-Gee, F., Bohnet, I., & Frey, B. S. (1997). Fairness and competence in democratic decisions. *Public Choice, 91*(1), 89–105.
Pampino, R. N., Jr, Heering, P. W., Wilder, D. A., Barton, C. G., & Burson, L. M. (2004). The use of the performance diagnostic checklist to guide intervention selection in an independently owned coffee shop. *Journal of Organizational Behavior Management, 23* (2–3), 5–19.
Patel, M. S., Volpp, K. G., Rosin, R., Bellamy, S. L., Small, D. S., Heuer, J., … Lee, S. M. (2018). A randomized, controlled trial of lottery-based financial incentives to increase physical activity among overweight and obese adults. *American Journal of Health Promotion, 32*(7), 1568–1575.
Paúl, M. L. (2022). *Raising Cane's drops $100,000 on Mega Millions tickets — again*. Retrieved 5 April 2023 from https://www.washingtonpost.com/nation/2022/07/29/raising-canes-mega-millions-lottery/.
PayScale. (2018). *2018 Compensation Best Practices Report - The Great Divide*. https://www.payscale.com/cbpr.
Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making, 29*(5), 453–469.
Pritchard, R. D., Leonard, D. W., Von Bergen Jr, C. W., & Kirk, R. J. (1976). The effects of varying schedules of reinforcement on human task performance. *Organizational Behavior and Human Performance, 16*(2), 205–230.
Raux, C., Souche, S., & Croissant, Y. (2009). How fair is pricing perceived to be? An empirical study. *Public Choice, 139*(1–2), 227–240.
Rosaz, J., & Villeval, M. C. (2012). Lies and biased evaluation: A real-effort experiment. *Journal of Economic Behavior & Organization, 84*(2), 537–549.
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(August)*, 556–570.
Rydval, O., Ortmann, A., Prokosheva, S., & Hertwig, R. (2009). How certain is the uncertainty effect? *Experimental Economics, 12*(4), 473–487.
Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology, 102*(3), 254.
Savage, D. A., & Torgler, B. (2010). Perceptions of fairness and allocation systems. *Economic Analysis and Policy, 40*(2), 229–248.
Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71*(3), 432.
Sen, A. P., Sewell, T. B., Riley, E. B., Stearman, B., Bellamy, S. L., Hu, M. F., … Loewenstein, G. (2014). Financial incentives for home-based health monitoring: A randomized controlled trial. *Journal of General Internal Medicine, 29*(5), 770–777.
Seybert, N., & Bloomfield, R. (2009). Contagion of wishful thinking in markets. *Management Science, 55*(5), 738–751.
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.
Simonsohn, U. (2009). Direct risk aversion: Evidence from risky prospects valued below their worst outcome. *Psychological Science, 20*(6), 686–692.
Soman, D., & Shi, M. (2003). Virtual progress: The effect of path characteristics on perceptions of progress and choice. *Management Science, 49*(9), 1229–1250.
Suff, P., Reilly, P., & Cox, A. (2007). Paying for Performance: New trends in performance-related pay. *Institute for Employment Studies*. https://www.employment-studies.co.uk/system/files/resources/files/mp78.pdf.
Tafkov, I. D. (2012). Private and public relative performance information under different compensation contracts. *The Accounting Review, 88*(1), 327–350.
van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J., Schipper, M., Zeelenberg, M., Berkhout, S., … Prast, H. M. (2018a). Commitment Lotteries Promote Physical Activity Among Overweight Adults—A Cluster Randomized Trial. *Annals of Behavioral Medicine, 52*(4), 342–351.
van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J., Schipper, M., Zeelenberg, M., Berkhout, S., … Prast, H. M. (2018b). Physical activity after commitment lotteries: Examining long-term results in a cluster randomized trial. *Journal of Behavioral Medicine*, 1–11.
Van Eerde, W., & Thierry, H. (1996). Vroom's expectancy models and work-related criteria: A meta-analysis. *Journal of Applied Psychology, 81*(5), 575–586.
Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: A randomized trial. *JAMA, 300*(22), 2631–2637.
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), 272.
Von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior* ((2nd ed.).). Princeton University Press.
Vroom, V. (1964). *Work and Motivation*. Wiley.
Wallace, S. G., & Etkin, J. (2018). How goal specificity shapes motivation: A reference points perspective. *Journal of Consumer Research, 44*(5), 1033–1051.
Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and Research*. Allyn and Bacon.
Weibel, A., Rost, K., & Osterloh, M. (2009). Pay for performance in the public sector—Benefits and (hidden) costs. *Journal of Public Administration Research and Theory, 20*(2), 387–412.
Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39*(5), 806–820.
Weiss-Cohen, L., Konstantinidis, E., Speekenbrink, M., & Harvey, N. (2016). Incorporating conflicting descriptions into decisions from experience. *Organizational Behavior and Human Decision Processes, 135*, 55–69.
Wiersma, U. J. (1992). The effects of extrinsic rewards in intrinsic motivation: A meta-analysis. *Journal of Occupational and Organizational Psychology, 65*(2), 101–114.
Wine, B., Edgerton, L., Inzana, E., & Newcomb, E. T. (2017). Further effects of lottery odds on responding. *Journal of Organizational Behavior Management, 37*(1), 75–82. https://doi.org/10.1080/01608061.2016.1267064
Wiseman, R. M., & Gomez-Mejia, L. R. (1998). A behavioral agency model of managerial risk taking. *Academy of Management Review, 23*(1), 133–153.
Wright, P. M. (1992). An examination of the relationships among monetary incentives, goal level, goal commitment, and performance. *Journal of Management, 18*(4), 677–693.
Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science, 42*(12), 1676–1690.

Zhang, Y., & Gao, L. (2016). Wanting Ever More: Acquisition Procedure Motivates Continued Reward Acquisition. *Journal of Consumer Research, 43*(2), 230–245. https://doi.org/10.1093/jcr/ucw017

Zuckerman, M. (2007). *Sensation Seeking and Risky Behavior*. American Psychological Association.

Zuriff, G. E. (1970). A comparison of variable-ratio and variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 13*(3), 369–374.