Journal of Experimental Psychology: General

Can Cognitive Discovery Be Incentivized With Money?

Pamela J. Osborn Popp, Ben R. Newell, Daniel M. Bartels, and Todd M. Gureckis

Online First Publication, January 20, 2025. https://dx.doi.org/10.1037/xge0001682

CITATION

Osborn Popp, P. J., Newell, B. R., Bartels, D. M., & Gureckis, T. M. (2025). Can cognitive discovery be incentivized with money? *Journal of Experimental Psychology: General*. Advance online publication. https://dx.doi.org/10.1037/xge0001682

Journal of Experimental Psychology: General

ISSN: 0096-3445
https://doi.org/10.1037/xge0001682

Can Cognitive Discovery Be Incentivized With Money?

Pamela J. Osborn Popp[1], Ben R. Newell[2], Daniel M. Bartels[3], and Todd M. Gureckis[1]
[1] Department of Psychology, New York University
[2] School of Psychology, University of New South Wales
[3] Booth School of Business, University of Chicago

The ability to discover patterns or rules from our experiences is critical to science, engineering, and art. In this article, we examine how much people's discovery of patterns can be incentivized by financial rewards. In particular, we investigate a classic category learning task for which the effect of financial incentives is unknown (Shepard et al., 1961). Across five experiments, we find no effect of incentive on rule discovery performance. However, in a sixth experiment requiring category recognition but not learning, we find a large effect of incentives on response time and a small effect on task performance. Participants appear to apply more effort in valuable contexts, but the effort is disproportionate with the performance improvement. Taken together, the results suggest that performance in tasks that require novel inductive insights is relatively immune to financial incentives, while tasks that require rote perseverance of a fixed strategy are more malleable.

Public Significance Statement

Cognitive discovery is the ability to abstract novel patterns to explain and interact with the world. Here, we explore whether the act of discovery can be promoted by increasing the financial stakes of a task. A popular idea in economics and psychology is that people devote additional cognitive resources as economic stakes increase. However, the limitations of this principle are unclear. We tested how rule discovery and rule following were modulated by financial incentives. While incentives did not improve rule discovery, we did observe that incentives improve rule following. The results give new insight into situations where incentives are likely (not) to improve cognitive performance, with implications for motivating behaviors in the workplace and in educational settings.

Keywords: learning, motivation, inductive reasoning, categorization, behavioral economics

Financial reward and incentives often motivate performance by increasing attention and effort to a task. However, less is known about the interaction of financial or other incentives with complex cognitive processes that require learning, insight, creativity, discovery, and induction. Can we incentivize these acts of creative cognitive discovery, and if not, why?

Under one broadly accepted view, people allocate attention and effort under an implicit cost–benefit analysis (Kool et al., 2017; Payne et al., 1993; Savage, 1972; Simon, 1990). Classic meta-analyses (Camerer & Hogarth, 1999) as well as more recent articles (Caplin et al., 2020; DellaVigna & Pope, 2018) find that attention and effort can be modulated by increasing financial incentives. For example, Caplin et al. (2020) showed that performance and the time participants spend on a simple perceptual task (e.g., counting the number of seven- vs. nine-sided polygons in a crowded display) scale with the financial reward value of each trial. Similar incentive and

Timothy Vickery served as action editor.
Pamela J. Osborn Popp https://orcid.org/0009-0007-5782-8895
All data, experiment code, and statistical analysis code are publicly available on Github: data and analysis code at https://github.com/NYUCCL/cognitive_effort and experiment code at https://github.com/NYUCCL/cognitive_effort_experiment. Results from the experiments presented here have been presented in talks at the Society for Mathematical Psychology, 2021, the Cognitive Science Society, 2022, and the 12th Triennial Invitational Choice Symposium, 2023. A conference article describing a subset of the experiments in this work was published in the proceedings of the 2022 Annual Meeting of the Cognitive Science Society.

The authors declare that they have no conflicts of interest. Todd M. Gureckis and Pamela J. Osborn Popp were supported by the National Science Foundation Grant 2121102. Ben R. Newell was supported by the Australian Research Council (Grant DP190101675). Daniel M. Bartels thanks the Kilts Center at the University of Chicago Booth School of Business for funding. The authors thank Andrew Caplin for helpful discussions on experiment design and participants at the 12th Triennial Choice Symposium for helpful feedback on this project.

Pamela J. Osborn Popp played a lead role in data curation, formal analysis, investigation, software, visualization, and writing–original draft and an equal role in conceptualization, methodology, and writing–review and editing. Ben R. Newell played an equal role in conceptualization, methodology, supervision, and writing–review and editing. Daniel M. Bartels played an equal role in conceptualization, methodology, supervision, and writing–review and editing. Todd M. Gureckis played a lead role in supervision and an equal role in conceptualization, methodology, and writing–review and editing.
Correspondence concerning this article should be addressed to Pamela J. Osborn Popp, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, United States. Email: pamop@nyu.edu

cost manipulations have been argued to modulate the trade-off between habitual and more cognitive forms of decision making (Kool et al., 2017).

However, one long-standing puzzle is that there are also behavioral tasks that appear to be surprisingly immune to the effects of incentive. Enke et al. (2021) found that incentives had only a weak effect on reducing cognitive biases in simple decision-making tasks but did increase response times (RTs). Similarly, van den Berg et al. (2023) found no effect of incentive on visual working memory performance. Additionally, the specific cognitive processes a task requires heavily influence its malleability in the face of incentives. For instance, one recent study found that financial incentives facilitated creativity in a constrained task, while creativity was unaffected by the same financial incentives in a more open-ended and unconstrained task (Charness & Grieco, 2019). Even though these tasks rely on elements of cognition like attention and effort, the reward modulation of higher order cognitive function is potentially more nuanced in part because of intrinsic constraints on human cognition in terms of time or computation (Simon, 1990). Even if the cost–benefit theory is correct, it is reasonable to expect to find certain tasks that require more cognitive computation than people can or will engage in (without resorting to extramental scaffolds such as using a calculator, computer, smartphone, or pen and paper). For example, we would not expect a large influence of incentives in tasks involving cognitive biases and heuristics for which performance is relatively unaffected by increased cognitive effort (see Stanovich & West, 2008, for examples of such tasks).

However, financial incentives are commonly used to reward intelligence (e.g., the Nobel Prize) or motivate learning (e.g., a parent giving their child $20 for every "A" grade they receive). These examples provide monetary incentives over a long timescale, as opposed to the short timescale of a psychology experiment. Still, it is possible that such incentives exist because of intrinsically held beliefs about the relationship between finances, effort, and cognition. It is important to identify if certain cognitive processes are outside voluntary control—you cannot punish a student for not understanding something if you acknowledge that the type of learning required is not dependent on individual effort. The current work sheds light on what aspects of cognition are influenced by financial motivation and how future research can identify the complex relationship between motivation, effort, and behavioral performance.

The present article concerns how category learning is modulated by incentives, given that it is a complex composite of many elementary cognitive functions. Category learning is ubiquitous in our mental lives (e.g., categorizing which foods do or do not give you a stomach ache). In many cases, people learn how to categorize various items by forming explicit hypotheses and testing their accuracy through trial and error. Finding patterns or rules that differentiate items is also a type of creative problem-solving required in many professions (e.g., "is this paper acceptable for publication or not?"). Successful category learning, particularly for categories composed of complex rules and patterns, requires coordination of several cognitive processes including sustained attention and effort but also working memory and creative, open-ended inductive reasoning, and even insight. Performance therefore depends on elements of cognition, which have shown differential sensitivities to incentives in isolation. While learning may benefit from financially motivated attention and effort, other cognitively complex requirements of the task could serve as a bottleneck for improving performance.

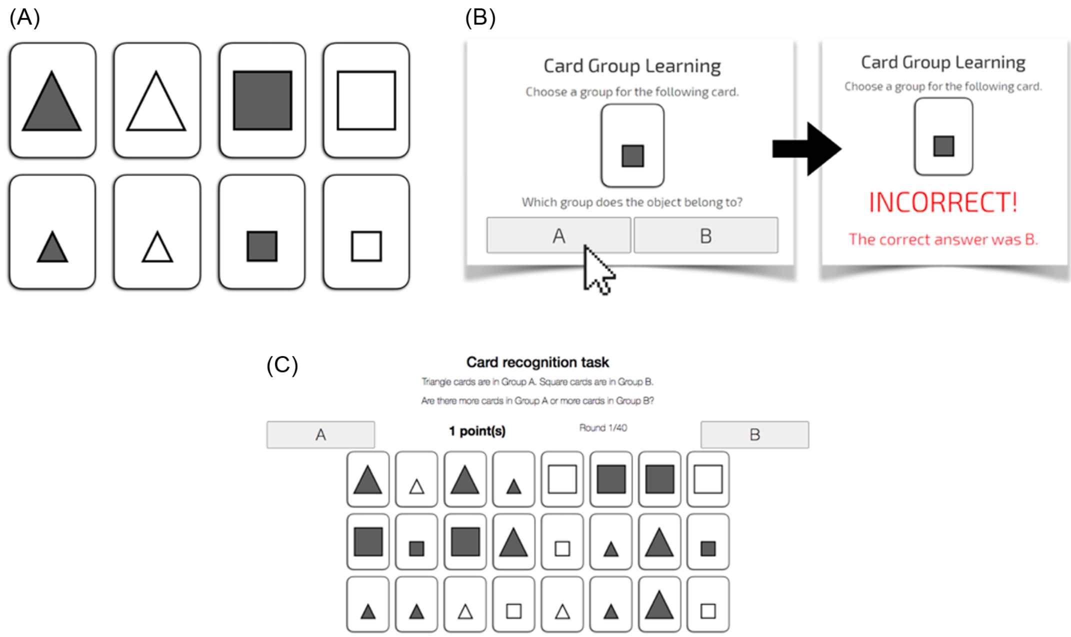
In the present study, we investigate the effect of financial incentive on performance in a well-studied category learning rule discovery task known as the SHJ task for the initials of the lead authors Shepard, Hovland, and Jenkins (Shepard et al., 1961). In the task, participants must attend to relevant stimulus features (e.g., the size, shape, or color of a simple shape) to correctly distinguish two groups of objects. For the same eight stimuli (see Figure 1), different classification rules serve to vary the difficulty of the task, as some categorizations are typically learned more quickly than others. Although participants could simply memorize the category memberships for the eight stimuli no matter the grouping (a challenging but by no means impossible task), a classic and highly replicated finding is that humans are generally biased to find an explanation of the categorization based on logical descriptions of stimulus features that define each category (Goodman et al., 2008; Kruschke, 1992; Love et al., 2004; Nosofsky et al., 1994).

One popular account proposes that participants approach the SHJ task by formulating and testing hypothetical rules; as they encounter new evidence trial-by-trial, they confirm (or reject and update) their hypotheses. Various models propose reasonable orders to hypothesis generation for this task, such as considering simpler rules before considering exceptions or more complex rules (Nosofsky et al., 1994). It has been established that participants narrow their focus and attention to relevant features (Rehder & Hoffman, 2005) as they settle on a hypothesis that produces no errors. Such accounts are consistent with the growing theoretical movement in cognitive science and economics that posits that behavior reflects a "resource-rational" trade-off between cognitive/computational costs and rewards (Bhui et al., 2021; Lieder & Griffiths, 2020). That is, participants apply only as much cognitive effort as required by the task (considering simple solutions before complex ones) and reduce cognitive load wherever possible (pruning hypotheses which have been proven false and exclusively attending to relevant stimulus feature dimensions).

By this logic, increasing task rewards would motivate participants to exert more cognitive effort, leading to improved performance depending on task difficulty. However, while processes like attention and menial effort can be financially incentivized (Caplin et al., 2020; DellaVigna & Pope, 2018), higher order cognitive processes vary in their sensitivity to incentivization. Work in the fields of insight and problem-solving has found that financial incentives promote performance in less creative tasks but weaken performance as tasks require more creative thinking to solve (Glucksberg, 1962). By some accounts, rule discovery can be seen as an "all-or-nothing" insight (Bower & Trabasso, 1963) since there is a marked point at which participants will attain the correct rule and make no further errors, meaning we could expect findings from the insight literature to apply here as well. However, rule discovery is notably different from the type of "Aha" moments that often arise during creative problem-solving tasks, since here, participants generate and incrementally update rules across trials in an evidence accumulation process that relies on memory. Ultimately, the SHJ task incorporates both low-level and high-level cognitive processes, and thus, it is not immediately apparent whether performance can be improved with incentives.

Based on considerable experience with this task and related variants, we began this project with the hypothesis that it may be difficult or

Figure 1
Experimental Task Stimuli and Appearance



Note. (A) Stimuli: The eight cards differing in three feature dimensions—shape, size, and color. Across different conditions, subjects learn through experience that half of the stimuli belong to the same group and are asked to discover the rule that determines group membership. (B) Category learning: An image from the learning phase of the category learning experiments (Experiments 1–5). In the learning phase, participants guess the category membership of each card and receive feedback on each trial. (C) Category recognition: The computer display for a single trial of the card category recognition task of Experiment 6. The top of the screen displays the number of points the participant will earn for a correct answer, and the rule distinguishing Group A cards from Group B cards. Participants are asked, "Are there more cards in Group A or more cards in Group B?" in regard to the 24 cards shown. See the online article for the color version of this figure.

impossible to motivate people to learn the SHJ categorization faster using various financial incentives. Prior work in this vein has provided mixed results. In a somewhat similar task, incentives did induce faster explicit rule learning under specific task conditions (Grimm et al., 2008; Maddox & Markman, 2010). On the other hand, a more recent article examining methods for obtaining high-quality data from Amazon Mechanical Turk (mTurk) found no effect of incentives on category learning performance in a subset of the SHJ categorization tasks (Crump et al., 2013). Importantly, none of these previous studies were designed or powered to address the question we raise. Our approach incorporates rigorous experimental economic methods to provide a novel account of whether financial incentives play a role in facilitating category learning and rule discovery.

Specifically, we aimed to perform a more comprehensive evaluation of incentives on category learning behavior that better controlled the rates and sizes of task payments inspired by recent work in economics on theories of rational inattention (Caplin et al., 2020). By varying the difficulty of the category learning task, we can gather more nuanced evidence of the interaction between difficulty and incentives. Additionally, we vary incentive systematically at different levels guided by both economic and psychological principles. In five category learning experiments and one category recognition experiment, we report the interactions and effects of difficulty and financial incentive on category discovery behavior. The results supported our initial intuitions and hypotheses and provide an important data point on the debate about resource-rational cognition. In particular, we show that it is, at the very least, difficult to financially incentivize the discovery or induction of a novel pattern or rule.

Method

We conducted six experiments on adult human participants to evaluate the effect of incentives on rule discovery and implementation. The first five experiments used a traditional category learning/rule discovery paradigm (Experiment 1) with variations to the number of learning trials (Experiment 2), the deterministic nature of the incentive (Experiment 3), within- versus between-subject task structure (Experiment 4), and preregistration and sample size (Experiment 5). The final experiment (Experiment 6) assessed the impact of incentives in a category recognition task that utilized the same stimuli and categorizations but did not require rule discovery.

Adult participants were recruited via Amazon mTurk for all six experiments, restricted to users in the United States who had a task approval rating of above 95%. The task was designed in JavaScript and delivered to the participants' browser via psiTurk (Gureckis et al., 2016). For most of the studies, data were collected until the sample size included at least 20 participants per between-subject condition. Experiments 5 and 6 were more highly powered, with Experiment 5 following a preregistered protocol where data collection continued until preset statistical tests yielded a conclusive result, or a preset budget had been depleted. Subjects received a base payment that corresponded with the expected length of completing the task at a rate of $0.15 per minute, and could receive a performance-based bonus of up to $10. To determine whether or not they would receive the bonus in experiments with stochastic rewards, participants would see their local clock on screen with time shown to the milliseconds. They would click on a button to stop the clock, and the last two digits of the stopped time (milliseconds) would be compared to their bonus probability; if the two digits were at or below the bonus probability, they would receive the bonus. A page in the instructions allowed subjects to test stopping and starting the clock to demonstrate that the last two digits were random and could not be controlled. We emphasized the random nature of the bonus following the procedure of Caplin et al. (2020) to minimize expectations of deception from the experimenters. Participants underwent a rigorous instructions phase followed by comprehension checks to ensure their complete understanding of the task and incentives before beginning the task. Although participants were asked not to take notes or pictures during the experiment, we also asked them to honestly report at the end if they had used any memory help, knowing that their payment would remain the same regardless of their response. In addition to these participants being excluded due to admission of using externalized memory aids, subjects were also excluded if they encountered an error in the experiment that prevented completion.

Stimuli

Figure 1, Panel A, displays the eight cards used in the experiment, each of which contains an object with a particular shape (square or triangle), size (large or small), and color (white or black). The eight stimuli can be divided into two equal groups, for a total of 70 unique ways (calculated with $\binom{8}{4}$). Each of these 70 groupings fits into one of six rule types described by Shepard et al. (1961), which differ in the number of stimulus features required to define the rule. For example, a Type I rule varies the group along a single feature dimension, giving a rule like "Large objects are in group A and small objects are in group B." A Type II rule groups the stimuli along two feature dimensions; for example, "Black triangles and white squares are in group A." Rule Types III, IV, and V rely on all three features and allow a "rule-plus-exception" type explanation such as "Large shapes are in category A, except for the white square." Rule Type VI applies to groupings that cannot be described by a simple feature-based rule. In these cases, the group membership of each stimulus must be memorized, making Type VI categorizations the hardest to learn.

Category Learning Task (Experiments 1–5)

Experiments 1–5 derive their design from the classic category learning task introduced by Shepard et al. (1961). During a learning phase, subjects use trial and error to actively learn the assignment of eight stimuli into two groups. Immediately after the learning phase, subjects perform a test (the test phase) in which they report the group membership of each stimulus once. The five experiments described below all retain this same basic structure but differ in the number of trials in the learning phase, the modality of the financial incentive, and whether the manipulations were between or within subjects.

Experiment 1 tested 418 subjects across 18 conditions, not including six subjects who admitted to using external help. The task took approximately 15 min, and subjects were paid a $2.25 base rate for their time, with the chance of earning a $10.00 bonus depending on their performance.

Experiment 1 tested all six rule types at three different incentive levels (18 conditions). Rule type was varied between subjects to avoid learning effects across blocks. In Experiment 1, the learning phase consisted of 16 trials (two repeats of each stimulus in a random order; Figure 1, Panel B). Participants were instructed that the purpose of the learning phase was simply to learn the groupings and did not determine their bonus. Their performance in the test phase determined the chance of winning the bonus.

The instructions explained how better performance on the test would increase their chance at winning a $10.00 bonus. To make the probabilistic nature of the incentive clear, we explained the probabilities in terms of pulling a marble out of a bag: "Imagine a bag full of red and blue marbles. We pull out a marble at random. If the marble is blue, you win the extra $10. If the marble is red, you receive only the base payment." Depending on the incentive condition the subject was assigned to, they would be shown a certain number of red marbles in the bag. Performance above chance on the test phase could turn some of the marbles in the bag blue, thus increasing their chance at winning the bonus. Since there were eight questions on the test, chance performance would mean getting four correct. Therefore, we replaced a red marble with a blue marble for each correct answer beyond chance performance, with a maximum possible four blue marbles to be earned if participants answered all eight questions correctly. If participants answered four or fewer of the eight test questions correctly, all of the marbles would remain red, meaning they had no chance to win the bonus.

The three incentive conditions varied the maximal probability of winning the bonus, which we represented by changing the total number of marbles in the bag. The low, medium, and high incentive conditions showed 64, eight, and four marbles in the bag, respectively. If a subject performed perfectly on the test and turned the maximum four marbles blue, this would correspond to a bonus probability of 6.25% in the low incentive condition, 50% in the medium incentive condition, and 100% in the high incentive condition.

Correspondingly, each correct answer above chance was worth approximately 1.6%, 12.5%, or 25% in the low, medium, and high incentive conditions, respectively. In Experiment 2, we collected 97 participants across four conditions, not including three subjects who admitted to using external help. Because the length of the learning phase had increased, we increased the base payment to $2.50. Experiment 2 replicated the design of Experiment 1 but with the number of learning trials increased to 32, so that participants saw four repeats of each of the eight stimuli during the learning phase. We gathered participants in four conditions: two rule type conditions (Type II and Type IV) crossed with two incentive conditions (low and high).

In Experiment 3, we collected 93 subjects across four conditions, not including seven subjects who admitted to using external help. As in Experiment 2, the base payment was $2.50, and subjects could earn a bonus of up to $0.64 if they were randomly assigned to the low incentive condition, or $10.00 in the high incentive condition. Since random/chance performance would generally get four of eight answers correct, subjects earned one ticket if their test score was 5, two tickets if their test score was 6, three tickets if their test score was 7, and four tickets if their test score was 8. In the low incentive condition, tickets were worth \$0.16 each—the expected value of the 1.6% chance at \$10 that correct answers above chance were worth in Experiments 1 and 2. Subjects assigned to the low incentive condition could earn a maximum bonus of \$0.64. In the high incentive condition, tickets were worth \$2.50 each, allowing a maximum bonus of \$10.00.

Experiment 4 had 31 subjects, each of whom completed four game blocks. This total does not include five subjects who admitted to using external help. The study took about 30 min to complete, and participants received a base rate payment of \$4.50 for their time. Participants were eligible to earn up to a \$10.00 bonus based on their performance. Participants took part in four consecutive games. In two of the games, participants could earn "blue bonus tickets" worth \$0.02 each. In the other two games, participants could earn "gold bonus tickets" worth \$1.23 each. Each game followed the procedure of the task in Experiment 3, such that participants could earn up to four tickets in each game depending on their performance in the test phase. This corresponded to maximal earnings of \$0.08 on the low incentive "blue ticket" games and \$4.92 on the high incentive "gold ticket" games.

Correct answers in high incentive games were therefore over 60 times more valuable than those in low incentive games. The order of the games was randomized for each subject. Experiment 5 collected data from 318 participants across three rule type conditions. This total did not include 50 subjects excluded due to admitting to using outside help and 18 subjects excluded due to experiment error. The design of Experiment 5 was almost identical to that of Experiment 4, in that each subject played two games for blue bonus tickets and two games for gold bonus tickets (four games total). However, where in Experiment 4 rule type was varied within subject, in Experiment 5, each subject was assigned to only one rule type condition and saw variations on the same rule type in all four games. Experiment 5 was preregistered (AsPredicted Preregistration No. 92503 can be found at https://aspredicted.org/T5X_VZM), and data collection continued until the allotted budget had been depleted.

Category Recognition Task (Experiment 6)

Experiment 6 collected data from 200 participants across three rule type conditions. This total did not include 10 subjects excluded due to admitting to using outside help. The experiment was designed very similarly to the polygon identification task in Caplin et al. (2020). Subjects performed 40 trials in which their task was to identify whether there were more Group A cards or Group B cards on the screen (Figure 1, Panel C). Before each trial, the text describing the rule was shown, and this text was also at the top of the screen during each trial. Each trial showed 24 cards that were sampled randomly from the eight stimuli such that 11 cards were in Group A and 13 cards were in Group B, or vice versa. Thus, the difference between the card group counts was kept constant at two for every trial.

Participants could earn up to 200 points total across all trials, with two trials offering 32 points for a correct answer, three trials offering 16 points, five trials offering 8 points, six trials offering 4 points, eight trials offering 2 points, eight trials offering 1 point, and eight trials offering 0 points. This distribution implemented the design of Caplin et al. (2020) to examine a spread of incentive values within each participant. Subjects were instructed that their chance at winning a \$10.00 bonus would be their final score minus 100, since chance performance would correspond with a score of 100 points. For example, if a subject scored 165 points, they had a 65% chance at winning the bonus. As in Caplin et al. (2020), subjects were not told how many points they had earned until the end of the experiment.

Rule type was varied between subjects, with each trial generating a randomized instance of that rule type. For instance, a participant assigned to the Type I condition might see a "small versus large" rule in one trial and "triangle versus square" rule in the next.

Modeling and Statistics

We use a Bayesian logistic regression model to predict the probability of a correct response as a function of incentive and rule type. Although not preregistered, we also included an interaction term between incentive and rule type to account for the possibility of rule types having differential sensitivity to incentive. The full logistic regression model is shown in Equation 1 below:

$$\log\left(\frac{\mathrm{Pr}}{1-\mathrm{Pr}}\right) = \beta_0 + \beta_1\mathrm{Inc} + \beta_2\mathrm{Rule} + \beta_3(\mathrm{Inc}\times\mathrm{Rule}) + \beta_4(\mathrm{Inc}|\mathrm{id}), \tag{1}$$

where Pr is the probability of a correct response on a given trial. The term Inc is a positive float variable for incentive, and the term Rule is a categorical variable encoding rule type. The term id refers to a given subject's identity. The precise model used for each experiment differed such that when the experiment design included within-subject manipulations, we included subject-specific factors on the within-variable.

The posterior values of the β parameters were estimated with Markov Chain Monte Carlo sampling in the Bambi python package (Capretto et al., 2022), with additional analysis and visualization using the ArviZ package (Kumar et al., 2019). The models used the default priors established by the Bambi package. To fit each model, four Markov Chain Monte Carlo chains each ran 2,000 samples, the first half of which were discarded as burn-in.

The Results section reports mean estimates $\bar{x}$ for coefficients on predictors of performance in the logistic regressions and median estimates *Mdn* and median absolute deviation (MAD) values for coefficients on predictors on linear regression RT models. We also report 95% highest density intervals (HDIs) for the relevant model parameters. Separate models were fit to data from the learning phase and test phase of the experiments. All the modeling results presented below are from models of the test phase unless otherwise noted.

Transparency and Openness

The data, experiment code, and statistical analysis code are publicly available on Github: data and analysis code at https://github.com/NYUCCL/cognitive_effort and experiment code at https://github.com/NYUCCL/cognitive_effort_experiment. Following procedure approved by the New York University Institutional Review Board (IRB-FY2016-231: Active Learning in Dynamic Environments), participants provided informed consent prior to participating in the online experiment. Data collection and storage were consistent with the guidelines enforced by the New York University Institutional Review Board.

Results

The Results section presents data and analyses from six individual experiments. A summary of the six experiment designs and their results can be found in Table 1.

In Experiments 1 through 5, we tested the effects of incentives on a task where subjects had a limited number of trials to learn a categorization rule using trial and error. The first four experiments provided slight variations on the task design to remove any incidental experimental design elements that could influence the result. These variations included increasing the length of the learning phase, varying the format of the incentive (whether performance increased the *probability* or the *magnitude* of the reward), and within- versus between-subjects designs. The fifth experiment was a high-powered preregistered replication of the within-subjects design used in Experiment 4 (details in the Method section). Across all five experiments, we evaluate the performance of subjects on a test after the learning phase and find little evidence that final test score performance was altered by the different incentive conditions. In contrast, we found robust effects across all experiments of category structure on performance (consistent with the original report of Shepard et al., 1961).

The sixth and final experiment was a variation on the procedure we call the "category recognition" task, where subjects identified the group membership of several stimuli with the categorization rule explicitly provided on the screen at all times. In contrast to all the other experiments, Experiment 6 did not require rule learning or discovery but rather only the cognitive implementation and execution of a rule. Unlike the patterns we observed in Experiments 1 through 5, task performance was positively modulated by increasing incentives in Experiment 6.

Statistical insights are based on the results of logistic regressions fit to the performance data and linear regressions fit to the RT data. Estimated regression coefficients are considered to confidently indicate positive (or negative) effects only when their 95% HDIs are entirely above (or below) zero with no overlap. In certain cases, coefficients whose HDIs do overlap zero are reported as possible weak or inconclusive effects.

The general discussion reflects on the differences between tasks to pinpoint the challenges in motivating rule discovery and category learning with financial incentives.

Experiment 1

Experiment 1 ($N = 418$) tested all six rule types considered by Shepard et al. (1961) at three different incentive levels for a total of 18 conditions. The design was completely between subjects to avoid learning effects across tasks. In Experiment 1, the learning phase consisted of 16 trials (two repeats of each stimulus in a random order). Participants were instructed that the purpose of the learning phase was simply to learn the groupings and did not determine their bonus. Instead, participants were told that their performance in the test phase determined the chance of winning the bonus.

Our results replicate the main effect of rule on performance from Shepard et al. (1961; Figure 2). A logistic regression with the six rule types and the incentive condition as predictors confirmed that Types II through VI had worse performance compared to Type I, with all five rule type coefficient means estimated to be negative and their upper tail HDIs falling beneath −0.1. Rule Type VI had the greatest effect on performance with a mean of −1.591, HDI [−2.162, −1.036]. As with all of the coefficients in these logistic regressions, we can interpret this value as the expected change in the log odds of a subject giving a correct response on the test phase in this condition as compared to the easiest, here rule Type I, condition. Across all incentive conditions and across learning and test trials, Type I was the easiest rule to learn and Type VI was the hardest rule to learn. The other four rules fell in between this range of difficulty.

However, we did not find a general positive effect of incentive on performance for any of the rules. The right panel in Figure 2 shows participants' performance for the three different incentive conditions across the six unique rule types. Although the expected value of perfect performance in the high incentive condition was over 15 times larger than that of the low incentive condition, performance remained roughly constant. We validated this qualitative observation by examining the interaction coefficients of rule type and incentive in the logistic regression model on performance. The fitted model estimated a zero effect of incentive on performance, with the incentive coefficients for all six rule types having their HDIs overlap 0. For instance, the effect of incentive in the Type VI condition had a mean of −0.013, HDI [−0.043, 0.018] (Figure 3).

Response times on the test trials for rule Types II and IV, to compare with the other experiments, are reported in Figure 4. A linear

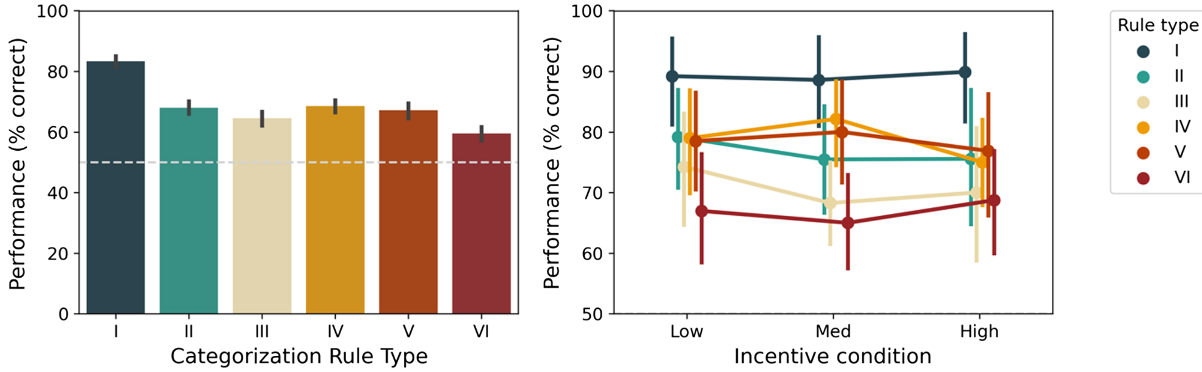
Table 1
Overview of Results for the Six Experiments

Experiment	Task	Design	*N*	Rule type	Incentive structure	Effect of incentive on performance?	Effect of incentive on response time?
1	Lng	B/S	418	I–VI	Prob	No	No
2	Lng	B/S	97	II, IV	Prob	No	No
3	Lng	B/S	93	II, IV	Mag	No	No
4	Lng	W/S	31	II, IV	Mag	No	No
5	Lng	W/S	318	I, II, IV	Mag	No	No
6	Rec	W/S	200	I, II, IV	Mag	Yes	Yes

Note. The first four columns distinguish the experimental design features, and the last two columns describe whether our statistical analyses found a meaningful effect of incentive on performance and response time, respectively. Task: Lng = category learning task; Rec = category recognition task. Design: B/S = between subjects; W/S = within subjects. Rule types: Which of the six rule type conditions were tested. Incentive structure: Prob = better performance increases *probability* of reward; Mag = better performance increases *magnitude* of reward.

Figure 2

Category Learning Performance in Experiment 1



Note. Left: Experiment 1 performance by rule type condition across all trials (both learning and test phases). Performance is percent correct across all learning and test phase trials. A gray dashed horizontal line shows performance at chance. Right: Experiment 1 test phase performance by incentive condition for the six rule types. Performance is measured only on test phase trials. Both: Error bars show 95% confidence intervals. See the online article for the color version of this figure.

regression model fit on the RT data showed an effect of rule type on RT but no effect of incentive on RT in the test phase (Figure 5). The coefficients on rule type, representing the intercept of the fitted regression line, were all above zero, generally increasing with the harder rule types. Meanwhile, the HDIs of the six rule type-incentive interaction coefficient estimates all overlap 0. A separate model fit to the learning phase data showed the same pattern of results with a smaller effect of rule type on RT.

Experiment 2

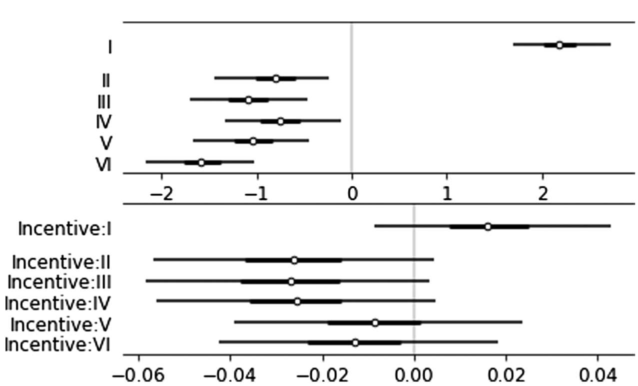
Given the null result of incentive in Experiment 1, we were concerned that 16 learning trials might have been inadequate for learning the categories. The number of learning trials had been selected to avoid ceiling effects on easier rule types based on learning curves from Nosofsky et al. (1994), but it may have created a floor effect on the more difficult rule type conditions. Therefore, we doubled the amount of learning trials for Experiment 2. In addition, we focused on rule Types II and IV, since these are both nontrivial but offer a range of difficulties. Rule Types II and IV are qualitatively different, and previous work has shown Type II performance to be better even when performance in Type III, IV, and V problems is indistinguishable (Nosofsky et al., 1994).

Experiment 2 ($N = 97$) replicated the design of Experiment 1 but with the number of learning trials increased to 32, so that participants saw four repeats of each of the eight stimuli during the learning phase. We assigned participants to four conditions: two rule type conditions (Type II and Type IV) crossed with two incentive conditions (low and high). As in the previous experiment, we saw little to no effect of incentive on performance in either of the rule conditions tested. The average performance by condition in Experiment 2 is shown in the top right plot in Figure 6. The logistic regression fit on the Experiment 2 data showed a small negative effect of incentive on performance in the rule Type II condition, $\bar{x} = -0.022$, HDI [−0.04, −0.005], indicating a very weak negative relationship between incentive and performance for this rule type. However, the coefficient's HDI overlaps with 0 and the effect is therefore not conclusive. The coefficient of incentive on performance in rule Type IV was slightly greater than that of rule Type II, $\bar{x} = 0.004$, HDI [−0.024, 0.032], but ultimately, the two rule type conditions did not have meaningfully different sensitivity to incentive.

There was a marginal increase in RT in the rule Type II test phase as a function of incentive (top right of Figure 4), with the regression model estimating the coefficient at a median of 24.54 (MAD = 8.799), HDI [−0.244, 49.16].

Figure 3

Experiment 1: Estimated Coefficients on Rule Type and Incentive in Logistic Regression Model of Performance



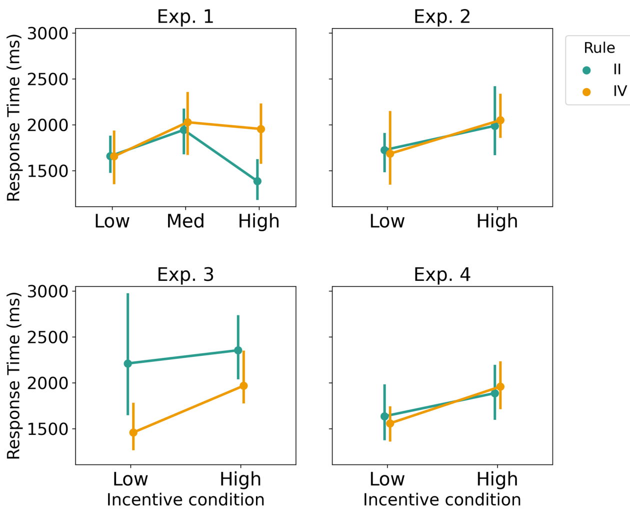
Note. The black lines reflect the 95% highest density interval, with the bold regions highlighting the interquartile range for each estimate. The Type I condition serves as the baseline regression line intercept from which the other conditions are compared. The interaction terms correspond to the effect of incentive in each condition, or the slope of the regression line; again, with values for Conditions II through VI serving as comparisons to the baseline of Condition I.

Experiment 3

Considering the null result from the first two experiments, we were concerned about the complexity of the probabilistic nature of

Figure 4

Median Response Times on Test Phase Trials Across Incentive Conditions for the First Four Category Learning Experiments



Note. See the online article for the color version of this figure.

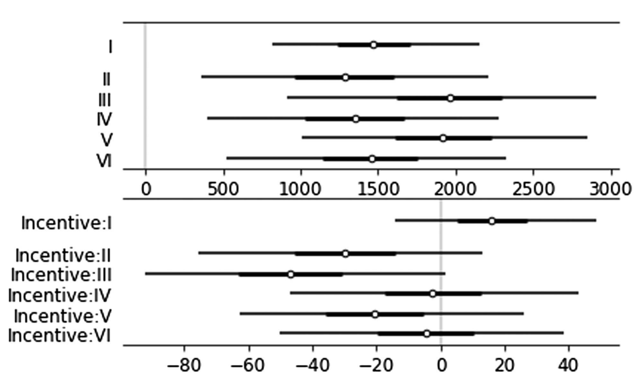
the incentive. The original probabilistic design was selected to build on the logic of other economic studies that intend for the likelihood of incentive to be interpreted linearly with its stated probability (see the Method section for details of the incentive structure based on Caplin et al., 2020). However, people's representation of probabilities may likely be nonlinear (Camerer & Ho, 1994; Tversky & Kahneman, 1992; Zhang et al., 2020), or participants may not have fully understood the probabilistic manipulation of incentive, or it might not have seemed meaningful. Not only is it challenging to convey such probabilistic information about the incentives, but it is also possible that the participants do not trust the legitimacy of the randomness determining their winnings.

The design of Experiment 3 ($N = 93$) was identical to Experiment 2, with the conditions tested being a cross of two rule types (Type II and IV) and two incentive levels (low and high). However, the incentive was not represented as a number of marbles in a bag but instead as the dollar value of tickets that participants would earn for their correct answers above chance on the test.

Even though there was a direct, certain relationship between participants' performance in the task and the magnitude of the bonus they would receive, performance still did not differ between the low and high incentive groups as shown in the bottom left plot in Figure 6. The model results showed the average effect of incentive on performance to be nearly zero with a mean of $\bar{x} = -0.019$, HDI [−0.038, 0.000]. There was also no effect of incentive on response time, as the small increase in RTs in the high incentive condition for rule Type IV was not robust (bottom left, Figure 4). The model fit on the RT data confirmed the lack of effect of incentive for both the rule Type II condition, HDI [−31.49, 35.66], and the rule Type IV condition, HDI [−23.82, 72.62].

Figure 5

Experiment 1: Estimated Coefficients on Rule Type and Incentive in Linear Regression Model of Response Time



Note. For more details, see Figure 3.

Experiment 4

Previously, we had avoided any within-subjects designs in order to prevent the learning effects that could come with performing multiple rounds of the category learning task. However, we were also concerned that modulating incentive between subjects possibly increased variance due to between-subject differences in sensitivity to small magnitude payments. If we were able to manipulate incentive within subject, a single participant would be able to see that incentive as a signal to devote more effort and attention to the more valuable portions of the experiment. The experiment that we used as a basis for our design, which had found effects of incentive on attention and effort, also manipulated incentive within subject (Caplin et al., 2020). As a result, we expected that if the SHJ task was sensitive to a voluntary change in cognitive approach, this experiment would finally reveal such an effect.

In Experiment 4 ($N = 31$), both rule type and incentive were varied within subject. The four game blocks each had a unique rule type and incentive pairing, crossing Types II and IV with a low and high incentive. As in Experiment 2 and Experiment 3, there were 32 learning trials in each game. Performance was incentivized by increasing magnitude rather than probability of reward, as in Experiment 3.

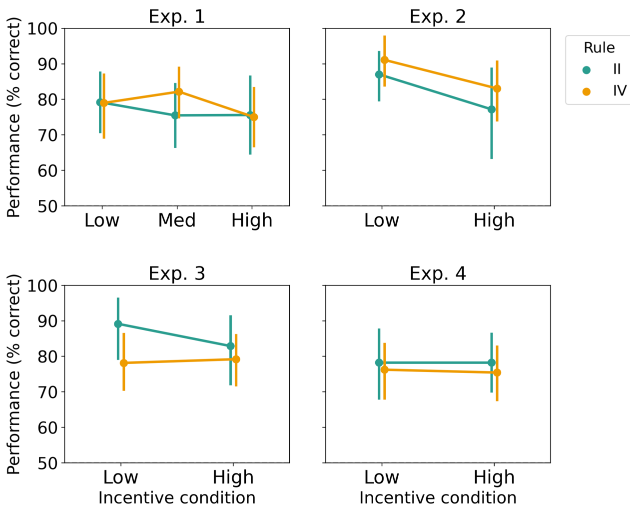
Although subjects now received more directly comparable signals of the relative value of each trial, we were surprised to see yet again no effect of the incentive manipulation on performance (bottom right, Figure 6). For the Experiment 4 analyses, the mixed effects model incorporated participant-level parameters to account for the within-subjects design. The null result was substantiated by the model fit on the data, showing no effect of incentive on performance for both rule Type II, HDI [−0.413, 0.429], and rule Type IV, HDI [−0.568, 0.527]. If subjects did apply more effort only on higher value portions of the experiment, this effort did not translate to better performance. There was no substantial increase in RT in the high incentive condition (bottom right, Figure 4), with the coefficient estimate overlapping zero in both rule type conditions.

Experiment 5

Experiment 5 is a preregistered, higher powered category learning experiment that closely resembled Experiment 4. In Experiment 5 ($N = 318$), rule type was varied between subjects, and three rule

Figure 6

Performance on the Test Phase Trials Across Conditions for the First Four Category Learning Experiments (Experiments 1–4)



Note. See the online article for the color version of this figure.

types were tested: Types I, II, and IV. As in Experiment 4, incentive was varied within subjects, with each subject performing two blocks of the task in which they were instructed that they were playing for low-value bonus tickets (worth \$0.02) and two blocks where they played for high-value bonus tickets (worth \$1.23).

The results showed no obvious effect of incentive on performance or response time (Figure 7). We again fit individual Bayesian mixed effect regression models to predict performance and response time as a function of rule type and incentive. The fitted logistic regression suggested that there was a small effect of incentive on performance but only in the Type II and Type IV conditions (Figure 8). The interquartile range of these coefficient estimates surpassed 0, but the HDI still overlapped 0, suggesting a very small, if any, effect. To evaluate the strength of the null hypothesis versus the alternative hypothesis, we calculated the Bayes factors (BFs) of frequentist paired t tests run on the three different rule type conditions. The t tests compared the null hypothesis that incentive had no impact on performance, with the alternative hypothesis that performance was greater in games with high-value bonus tickets. The BFs then compare the strength of evidence for the null hypothesis over the alternative, BF_{01}. For Type I, $BF_{01}[I] = 4.98$, with a t statistic of $t(223) = 0.77$, $p = .78$. For Type II, $BF_{01}[II] = 1.60$, $t(211) = -1.69$, $p = .046$, and finally for Type IV, $BF_{01}[IV] = 2.98$, $t(199) = -1.24$, $p = .11$. All of the BF_{01}s were above 1, suggesting stronger evidence for the null than the alternative hypothesis.

Experiment 6

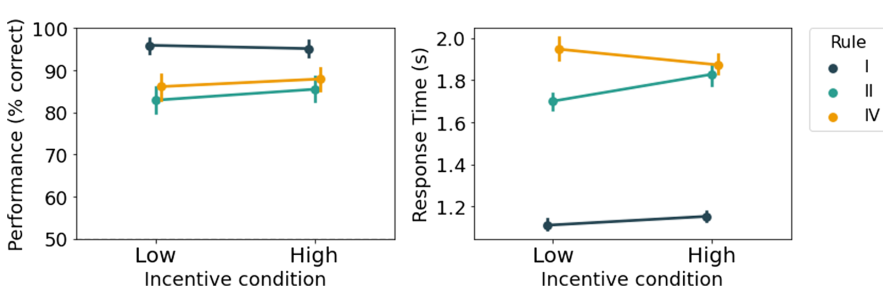
After this streak of null results for incentive (with multiple robust replications of the original categorization effects from SHJ), we wanted to restrict our design even further to be as similar as possible to studies that found effects of incentive on performance, such as Caplin et al. (2020). Our results so far suggested that rule inference and category learning are not influenced by incentive, in spite of many experiment variations. If we removed the requirement for subjects to execute the complex strategy of inductive inference, would we then see an effect of incentive on performance?

Therefore, Experiment 6 ($N = 200$) consisted of a task that relied on simply using a provided rule to categorize cards and determine whether there were more instances of cards in Group A or cards in Group B on the screen on a trial. The task would be more difficult for more complex rules, so we expected to see an effect of rule type on performance as in the previous experiments. However, the relevant rule would be provided to participants, removing their reliance on inference and memory.

Figure 9 shows performance increasing with the point value of the trials. As with the other experiments, we fit a logistic regression model to the performance data and a separate linear model to the response time data. The models contained subject-level predictors to account for random within-subjects effects of incentives. In the model predicting performance, the estimate of the coefficient on incentive in rule Type I was greater than zero, with a mean of $\bar{x} = 0.022$ ($SD = 0.009$, HDI [0.006, 0.039]). The incentive interaction with the rule Type II condition added additional positive slope with a coefficient of $\bar{x} = 0.009$ ($SD = 0.012$, HDI [−0.013, 0.032]). The effect of incentive on performance in the rule Type IV condition was approximately the same as the baseline of Type I, $\bar{x} = 0.000$ ($SD = 0.011$, HDI [−0.022, 0.021]). From this result, we conclude that incentive has a positive effect on performance. A visualization of these parameters is shown in Figure 10.

Figure 7

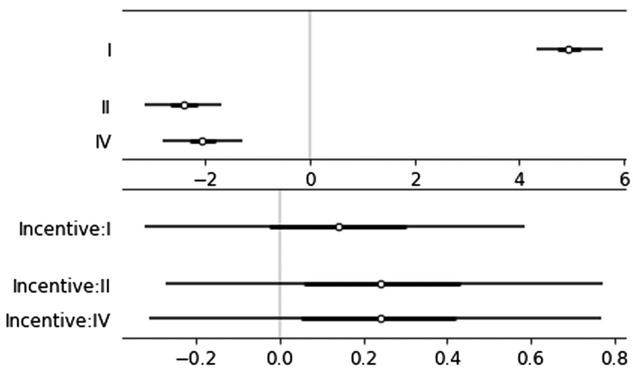
Category Learning Performance and Response Time in Experiment 5



Note. Left: Experiment 5 mean performance by incentive condition. Expected chance performance is 50%. Right: Experiment 5 median response times by incentive for the three rule types. See the online article for the color version of this figure.

Figure 8

Experiment 5: Estimated Coefficients on Rule Type and Incentive in Logistic Regression Model of Performance



In Experiment 6, we also found substantial effects of incentive and rule type on RT. The baseline incentive parameter in the regression model fit on RT had a median value of 80.97 (MAD = 58.1), HDI [−78.2, 247.8]. As we can see in the bottom part of Figure 11, the effect of incentive on response time was even larger in the rule Type II condition and largest in the Type IV condition. There was also a consistent effect of rule type alone on response time. Compared to the rule Type I condition, participants spent longer on each trial in the rule Type II condition (*Mdn* = 4,437, MAD = 1,704, HDI [−372.9, 9349]) and rule Type IV condition (*Mdn* = 7,110, MAD = 1,722, HDI [1369, 1.167×10^4]).

General Discussion

Across six experiments, we investigated the effects of financial incentives on category learning and recognition in the classic SHJ categorization task (Shepard et al., 1961). Each experiment followed up on questions unresolved in preceding ones and included replications of key results across studies. We found a substantial (nonzero) positive effect of incentive on performance only in a task where rule implementation (Experiment 6), and not rule discovery (Experiments 1–5), was required. In addition, when an effect of incentive on performance was found, we also found an effect of incentive on response time. Overall, the results suggest that when a task requires sustained and careful effort, rather than learning and insight, incentives can motivate better performance.

The key question then is why rule discovery seems resistant to incentives. Category rule discovery learning is a costly cognitive operation because it requires forming and testing alternative hypotheses. Still, incentives could encourage participants to hold multiple hypotheses in mind to test at the same time, allowing for a more meaningful change to the hypothesis space when new feedback is received. Incentives could also inspire participants to memorize the category label for each stimulus, in which case they would only need one trial per stimulus to produce maximum performance. However, our results reveal that participants do not use either of these approaches, perhaps because they are not apparent or not physically feasible.

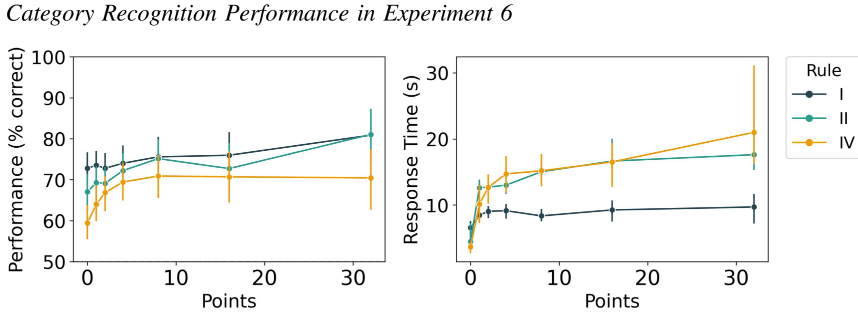
The financial incentives used here do not reduce the space of possible hypotheses, provide additional information, suggest alternate strategies, or give specific direction that could make the task "easier" (less cognitively effortful). This was central to our initial hypothesis about the resistance of category rule discovery to incentive modulation.

However, incentives might elevate performance if they can improve participants' ability to attend to and remember the examples they had viewed recently and integrate across them to abstract a pattern. Monetary incentives have been found to modulate top-down attention in both behavioral (Caplin et al., 2020) and neural data (Padmala & Pessoa, 2011; Small et al., 2005). Additionally, previous work has shown that people will incur monetary costs to avoid implementing complex rules. Oprea (2020) found that people have a preference for cognitively simple rules, and one might predict that financial incentives would encourage better performance on complex rules in the category learning task.

However, the key distinction seems to result from rule implementation versus rule learning. Integration across examples, as is required for rule learning, presumably requires working memory (Maddox & Markman, 2010), which is known to have limited capacity (Cowan, 2010) that may not be improved by incentives (Heitz et al., 2008; van den Berg et al., 2023). Perhaps because rule discovery relies on both attention and working memory, it remains resistant to monetary incentives. While the nature of rule discovery as a process of evidence

Figure 9

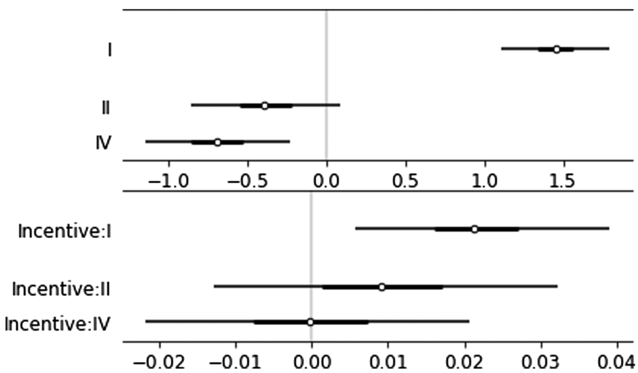
Category Recognition Performance in Experiment 6



Note. Left: Experiment 6 mean performance by incentive condition—the number of points a trial is worth—for the three rule types tested. A gray dashed horizontal line at 50% shows expected chance performance. Right: Experiment 6 median response times by incentive for the three rule types tested. See the online article for the color version of this figure.

Figure 10

Experiment 6: Estimated Coefficients on Rule Type and Incentive in Logistic Regression Model of Performance



Note. For more details, see Figure 3.

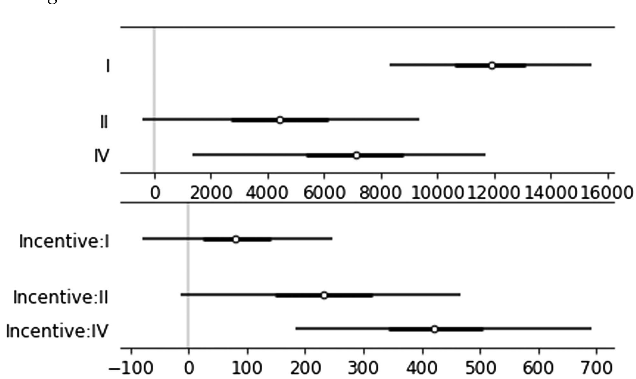
accumulation across trials distinguishes it from other forms of insight and problem-solving (as discussed in the introduction), its resistance to incentives suggests it may have more in common with tasks requiring creativity that similarly do not benefit from incentives.

At the same time, the results are somewhat unexpected based on common views about how cognitive computations trade off against utility. Economic labor supply theory proposes a utility function that balances time spent on effort, assuming effort yields income, and leisure, which does not yield income (Botvinick & Braver, 2015; Kool & Botvinick, 2014). The prediction that naturally follows is that when large financial incentives are tied to task performance, participants will allocate more time to that task than when incentives are smaller. Indeed, we see such an effect in Experiment 6, in which the need to search through a hypothesis space of possible rules was removed: Both participants' response time and performance increased.

However, similar financial incentives applied to a rule discovery task did not yield an increase in response time or an increase in performance, which suggests that task costs in the form of cognitive limitations cannot always be overcome by effort and motivation. This finding may apply in particular to tasks that rely heavily on insight, discovery, and creativity. We connect this result to the conclusions of Charness and Grieco (2019), who found that financial incentives augmented creativity in constrained but not open-ended task environments. The rule discovery and rule implementation tasks used here similarly require differing amounts of creative thought (or rather exploration of hypotheses), which could explain their different sensitivity to reward.

Figure 11

Experiment 6: Estimated Coefficients on Rule Type and Incentive in Linear Regression Model of Response Time in the Category Recognition Task



In the same vein, an alternate explanation of our results is that the rule discovery task is not completely immune to incentives, but instead has a much higher monetary threshold because of the costliness of the cognition required (e.g., hundreds rather than tens of dollars in reward). In our rule discovery experiments, performance in high incentive conditions or trials was as little as 15 times (Experiment 1) and as much as 60 times (Experiments 4 and 5) more valuable than in the low incentive condition. While the absolute and relative values of the rewards were enough to encourage a performance increase in the rule implementation task, they did not have the same effect on rule discovery. That a larger financial incentive could yield a rule discovery performance effect does not run counter to our conclusions but is not a question we can answer with the data collected here.

Another potentially relevant framing is that of the dual-process theory of cognition, which distinguishes two main systems that are claimed to direct judgment and decision making (Dolan & Dayan, 2013; Evans, 2008; Kahneman & Frederick, 2002). "Type 1" processing is construed as largely automatic and intuitive, while "Type 2" processing is said to consist of analytical, effortful consideration. The rule discovery task described here could be described in terms of the combination of such automatic and analytic processes. For example, participants may memorize labels and simply access the memory when they choose a group category in a more associative "Type 1" manner. They can also test and use specific rules from a hypothesis space in a more deliberate "Type 2" fashion. The same is true for the rule implementation task; while providing the rule to participants may reduce their need to strategically consider multiple hypotheses, the task still necessitates effortful attention to apply and count instances of a rule. Labeling these different strategies—memorization and rule testing—as indicative of Type 1 and Type 2 processing, respectively, is common but, to our minds, does not add explanatory value in this particular instance. While we cannot directly rule out a dual-process account with our data (in part because of the inherent flexibility of such theories; Newell & Shanks, 2023), we leave the pursuit of experimental designs that could provide unequivocal support for such an interpretation to future research.

It is also important to mention that our approach exclusively examined financial incentives. However, it is known that both financial and nonfinancial incentives can have varying impacts on effort and performance (Bonner & Sprinkle, 2002; Read, 2005). In the creativity experiment mentioned above (Charness & Grieco, 2019), a social motivation (ranking among one's peers) facilitated creativity in both the constrained and unconstrained tasks. While we did not use any nonmonetary incentives in the present study, it is unclear whether rule discovery performance is sensitive to motivations of any format, financial or otherwise. Future work should investigate the mismatch between financial and social incentives' effects on creative versus rote tasks.

While the current article did not find strong effects of incentives on category learning, we considered several possible ways that financial incentives could be incorporated into existing models of categorization performance (e.g., Turner, 2019). For instance, incentives might lead to lower lapse rates, higher learning rates, or simpler strategies. Our preliminary investigations into whether participants were more likely to learn easier rules in low incentive conditions were inconclusive, owing to the fact that the experiment was designed to highlight performance and not strategy. A promising future direction would be to use a task that includes transfer stimuli in order to more directly assess the specific rules participants are learning.

The current results suggest that a comprehensive framework of motivation and performance needs to acknowledge that cognitive processes have different malleability to incentives. Rule discovery appears to be relatively impervious to incentives because it requires flexible cognition that depends on creativity and limited resources such as working memory. In contrast, rote tasks like rule recognition rely on simple cognitive routines that can be easily compiled and repeated and, as a result, are impacted positively by increasing incentives. The extent to which this conclusion holds for other varieties of tasks that lie on a continuum from discovery to implementation is an important focus for future work. In addition, mapping out the impact of incentives on performance as people transition from the discovery to the implementation of a rule will likely yield further insight into the complex interplay between cognitive effort, cognitive limitations, and motivation.

Constraints on Generality

Participants that took part in the experiments were adults on Amazon mTurk who were registered as located in the United States, English-speaking, and who had a task approval rating of above 95%. Participants thereby necessarily had a basic amount of computer literacy as required to participate in online crowdsourced psychology research. Participants were only excluded from the analysis if they admitted to using external memory aids to remember the stimuli category assignments. An optional, brief self-reported demographic survey at the end of the study indicated that the mean age was 38 years old (SD = 10), 62% of participants identified as male, 37% identified as female, and over 91% of participants had at least some posthigh school education. We believe the results are likely to reflect trends among the general English-speaking population, with potentially less application to gender minorities or to those with less than a high school education.

References

Bhui, R., Lai, L., & Gershman, S. J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, *41*, 15–21. https://doi.org/10.1016/j.cobeha.2021.02.015

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. https://doi.org/10.1016/S0361-3682(01)00052-6

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*(1), 83–113. https://doi.org/10.1146/annurev-psych-010814-015044

Bower, G., & Trabasso, T. (1963). Reversals prior to solution in concept identification. *Journal of Experimental Psychology*, *66*(4), 409–418. https://doi.org/10.1037/h0044972

Camerer, C. F., & Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, *8*(2), 167–196. https://doi.org/10.1007/BF01065371

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), 7–42. https://doi.org/10.1023/A:1007850605129

Caplin, A., Csaba, D., Leahy, J., & Nov, O. (2020). Rational inattention, competitive supply, and psychometrics. *The Quarterly Journal of Economics*, *135*(3), 1681–1724. https://doi.org/10.1093/qje/qjaa011

Capretto, T., Piho, C., Kumar, R., Westfall, J., Yarkoni, T., & Martin, O. A. (2022). Bambi: A simple interface for fitting Bayesian linear models in python. *Journal of Statistical Software*, *103*(15), 1–29. https://doi.org/10.18637/jss.v103.i15

Charness, G., & Grieco, D. (2019). Creativity and incentives. *Journal of the European Economic Association*, *17*(2), 454–496. https://doi.org/10.1093/jeea/jvx055

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLOS ONE*, *8*(3), Article e57410. https://doi.org/10.1371/journal.pone.0057410

DellaVigna, S., & Pope, D. (2018). What motivates effort? Evidence and expert forecasts. *The Review of Economic Studies*, *85*(2), 1029–1069. https://doi.org/10.1093/restud/rdx033

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Enke, B., Gneezy, U., Hall, B., Martin, D. C., Nelidov, V., Offerman, T., & van de Ven, J. (2021). *Cognitive biases: Mistakes or missing stakes?* (Tech. rep.). National Bureau of Economic Research. https://doi.org/10.3386/w28650

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Glucksberg, S. (1962). The influence of strength of drive on functional fixedness and perceptual recognition. *Journal of Experimental Psychology*, *63*(1), 36–41. https://doi.org/10.1037/h0044683

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive Science*, *32*(1), 108–154. https://doi.org/10.1080/03640210701802071

Grimm, L. R., Markman, A. B., Maddox, W. T., & Baldwin, G. C. (2008). Differential effects of regulatory fit on category learning. *Journal of Experimental Social Psychology*, *44*(3), 920–927. https://doi.org/10.1016/j.jesp.2007.10.010

Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., Halpern, D., Hamrick, J. B., & Chan, P. (2016). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, *48*(3), 829–842. https://doi.org/10.3758/s13428-015-0642-8

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, *45*(1), 119–129. https://doi.org/10.1111/j.1469-8986.2007.00605.x

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.004

Kool, W., & Botvinick, M. (2014). *A labor/leisure tradeoff in cognitive control*. Educational Publishing Foundation.

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost–benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, *28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44. https://doi.org/10.1037/0033-295X.99.1.22

Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. (2019). Arviz a unified library for exploratory analysis of Bayesian models in python. *Journal of Open Source Software*, *4*(33), Article 1143. https://doi.org/10.21105/joss.01143

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, Article e1. https://doi.org/10.1017/S0140525X1900061X

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*(2), 309–332. https://doi.org/10.1037/0033-295X.111.2.309

Maddox, W. T., & Markman, A. B. (2010). The motivation–cognition interface in learning and decision making. *Current Directions in Psychological Science*, *19*(2), 106–110. https://doi.org/10.1177/0963721410364008

Newell, B. R., & Shanks, D. R. (2023). *Open minded: Searching for truth about the unconscious mind*. MIT Press. https://doi.org/10.7551/mitpress/14922.001.0001

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*(1), 53–79. https://doi.org/10.1037/0033-295X.101.1.53

Oprea, R. (2020). What makes a rule complex? *American Economic Review*, *110*(12), 3913–3951. https://doi.org/10.1257/aer.20191717

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, *23*(11), 3419–3432. https://doi.org/10.1162/jocn_a_00011

Payne, J. W., Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173933

Read, D. (2005). Monetary incentives, what are they good for? *Journal of Economic Methodology*, *12*(2), 265–276. https://doi.org/10.1080/13501780500086180

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*(1), 1–41. https://doi.org/10.1016/j.cogpsych.2004.11.001

Savage, L. J. (1972). *The foundations of statistics*. Courier Corporation.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, *75*(13), 1–42. https://doi.org/10.1037/h0093825

Simon, H. A. (1990). Bounded rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Utility and probability* (pp. 15–18). Palgrave Macmillan. https://doi.org/10.1007/978-1-349-20568-4_5

Small, D. M., Gitelman, D., Simmons, K., Bloise, S. M., Parrish, T., & Mesulam, M.-M. (2005). Monetary incentives enhance processing in brain regions mediating top-down control of attention. *Cerebral Cortex*, *15*(12), 1855–1865. https://doi.org/10.1093/cercor/bhi063

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*(4), 672–695. https://doi.org/10.1037/0022-3514.94.4.672

Turner, B. M. (2019). Toward a common representational framework for adaptation. *Psychological Review*, *126*(5), 660–692. https://doi.org/10.1037/rev0000148

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/BF00122574

van den Berg, R., Zou, Q., Li, Y., & Ma, W. J. (2023). No effect of monetary reward in a visual working memory task. *PLOS ONE*, *18*(1), Article e0280257. https://doi.org/10.1371/journal.pone.0280257

Zhang, H., Ren, X., & Maloney, L. T. (2020). The bounded rationality of probability distortion. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(36), 22024–22034. https://doi.org/10.1073/pnas.1922401117

Received April 3, 2023
Revision received August 22, 2024
Accepted September 24, 2024 ■