


EMPIRICAL ARTICLE

# Do losses trigger deliberative reasoning?

Jeffrey Carpenter<sup>1</sup> and David Munro<sup>2</sup> 

<sup>1</sup>IZA and Department of Economics Middlebury College, Middlebury, VT, USA and <sup>2</sup>Department of Economics, Middlebury College, Middlebury, VT, USA

**Corresponding author:** David Munro; Email: [dmunro@middlebury.edu](mailto:dmunro@middlebury.edu)

**Received:** 25 May 2023; **Revised:** 17 April 2024; **Accepted:** 9 May 2024

**Keywords:** dual process theory; cognitive effort; bias; incentives; loss; experiment

## Abstract

There is a large literature evaluating the dual process model of cognition, including the biases and heuristics it implies. However, our understanding of what causes effortful thinking remains incomplete. To advance this literature, we focus on what triggers decision-makers to switch from the intuitive process (System 1) to the more deliberative process (System 2). We examine how the framing of incentives (gains versus losses) influences decision processing. To evaluate this, we design experiments based on a task developed to distinguish between intuitive and deliberative thinking. Replicating previous research, we find that losses elicit more cognitive effort. Most importantly, we also find that losses differentially reduce the incidence of intuitive answers, consistent with triggering a shift between these modes of cognition. We find substantial heterogeneity in these effects, with young men being much more responsive to the loss framing. To complement these findings, we provide robustness tests of our results using aggregated data, the imposition of a constraint to hinder the activation of System 2, and an analysis of incorrect, but unintuitive, answers to inform hybrid models of choice.

## 1. Introduction

It is well documented that individuals make systematic errors in various decision-making contexts, including specific areas such as finance (Frydman and Camerer, 2016; Malmendier, 2018; Thaler, 2005), organization (Camerer and Malmendier, 2007) or even accounting (Siegel and Ramanauskas-Marconi, 1989). They also suffer bias when making more general choices such as forecasting (Dong et al., 2021), forming beliefs (Alós-Ferrer and Garagnani, 2023) and assessing risk (Shepherd et al., 2015). Because these biases can lead to sub-optimal outcomes, utilizing choice architecture to minimize their effects has been demonstrated to improve decision-making and employee performance in some situations (e.g., Hossain and List, 2012; Huang et al., 2019).

A prominent conceptualization of biased decision making is the “dual process” or “dual system” framework first mentioned by James (1890) and popularized more recently in Kahneman (2011). In this framework, System 1 controls our “gut” responses and acts quickly, requiring limited cognitive effort. For less obvious or intuitive choices, relying too heavily on System 1 can lead to biased decisions. By contrast, System 2 represents more effortful, contemplative cognition, and consumes more time, energy, and attention, but tends to drive us toward better choices. This framework forms the foundation for much of the behavioral research in social sciences, and recent evidence suggests many consumer biases

can be accounted for with processing-based cognitive models (Stango and Zinman, 2023). However, our understanding of what causes effortful thinking is incomplete (e.g., Westbrook and Braver, 2015). A particularly relevant question relating to the dual system framework (highlighted in the recent survey article of De Neys, 2023) involves understanding what causes individuals to switch from System 1 (S1) to System 2 (S2), since this distinction is thought to be at the core of many flawed choices. Of course, it is important to emphasize at the outset that S1 and S2 do not represent any particular parts of the brain; the typology should not be considered exhaustive of all modes of cognition; nor should these systems be thought of as necessarily mutually exclusive (e.g., Melnikoff and Bargh, 2018). They are simply coarse metaphors for heuristic versus more deliberative forms of thinking. Given the important role these different forms of thinking play in understanding decision making, our focus is to better understand what causes individuals to engage in more or less heuristic versus deliberative thinking.

In particular, we explore how the context of incentives, gains versus losses, marshal cognitive resources, influence the propensity for S1 and S2 thinking and, ultimately, impact decision making.<sup>1</sup> In an online experiment with three incentive treatments, we test performance on seven Cognitive Reflection Test (CRT) questions, designed to differentiate between S1 and S2 thinking by having an intuitive, but incorrect, answer. To formulate hypotheses about what economically-relevant factors might trigger deliberative S2 reasoning, we rely on a literature in economics, cognitive psychology and neuroscience showing that the framing of rewards can affect the amount of cognitive effort people devote to decisions. Our motivation stems from the idea that cognitive resources are scarce and decision-makers must economize. In doing so, the incentives related to employing cognitive effort will naturally influence whether a decision-maker chooses to use scarce resources on a task. In switching between systems, there is undoubtedly a threshold beyond which a stimulus must cross to activate the more costly S2. Perhaps a useful analogy from neuroscience is an action potential, where a specific threshold voltage is required to send neuronal signals. In many economic contexts where changing behavior is costly, there is a threshold of incentives required to elicit a response, otherwise the decision-maker remains inactive. Our interest is in the differential perceptions of gains and losses and how this framing of rewards influences the propensity to engage in S2 thinking. There are multiple interpretations as to why losses might differentially stimulate S2 thinking. One interpretation is that if the subjective weight given to losses is greater (i.e., loss aversion), the incentive stimulus of losses will be greater and may elicit more effort. We discuss the literature on losses and effort, in general, below. An alternative interpretation is that, regardless of the preferences over monetary gains/losses, losses are stronger attractors of attention. For example, in the loss-attention model of Yechiam and Hochman (2013), losses increase on-task attention and enhance the sensitivity to the reinforcement structure even in the absence of loss aversion. Both of these mechanisms highlight the possibility of a differential effect of losses on decision making, which motivates our use of these frames to understand the drivers S1 and S2 thinking.

While numerous models of human cognition exist, many focus on how individuals arrive at correct solutions to problems (e.g., the “Aha!” model, Kounios and Beeman, 2009). We are interested in models that distinguish between the types of incorrect answers individuals give. Dual process models are somewhat unique in that they posit that, in addition to correct and incorrect answers, there is a third class of answer that is particularly interesting to study: intuitive responses that seem obvious and correct but are not. It is this third type of response that identifies a low-effort heuristic mode of reasoning (S1). Further, the standard approach used to evaluate the hypothesis that cognition is regulated by these two processes is to use CRT questions to distinguish between intuitive versus contemplative reasoning (Frederick, 2005). Notice, however, that this paradigm also permits participants to provide incorrect answers that are not intuitive. This allows us to study not just how incentive framing causes individuals to arrive at correct solutions, it also allows us to explore the potential for multiple decision processes more broadly.

---

<sup>1</sup>The association between threats, vigilance and the propensity for S2 thinking is an important postulate in Kahneman (2011), and our paper can be viewed as an exploration of those general ideas.

There are two recent meta-analyses on CRT performance. Brañas-Garza et al. (2019) examine 118 CRT studies and find, among other things, that monetary incentives do not impact performance. However, in a meta-analysis examining studies that directly compared incentive versus no-incentive conditions, Yechiam and Zeif (2023) find evidence of a small positive effect on incentivization. Our results below find support for the latter conclusion, but that the framing of incentives also matters. Of interest, most studies utilizing CRT questions are simply interested in CRT performance (frequency of correct answers) and the various correlates with that. For example, in the meta-analysis of Brañas-Garza et al. (2019) there are no results presented on the frequency of intuitive answers. This is a missed opportunity to develop a deeper understand of the causes/correlates of heuristic versus deliberative thinking. By examining whether incentives mediate the progression from intuitive but incorrect answers to demonstrably incorrect answers given after some effort, and finally to correct ones, we can gain a deeper understanding of how cognitive effort relates to successful problem solving and what this says about different modes (and models) of cognition.

The existing literature has shown a general effect of incentives and rewards on cognitive effort. For example, Smith and Walker (1993) show that across 31 studies, increased rewards can shift choices toward the predictions of rational models, an outcome consistent with overcoming the opportunity cost of providing more cognitive effort. More recently, Clay et al. (2017) focus on how decision-makers demonstrate increased vigilance and attention when losses are possible, Lejarraga et al. (2019) show that losses prompt more attention and cognitive effort than gains, and Chen et al. (2020) extend these results showing that experimental participants expend disproportionately more effort to avoid losses even when the task is mostly physical.<sup>2</sup> There is a renaissance of work examining issues related to bounded rationality, and in particular how complexity and attention can influence decision making.<sup>3</sup> Important questions in this literature involve how the complexity of problems (e.g., whether they can be solved using simple rules or not) and limited attention (i.e., whether people understand the constraints on their cognitive resources and react optimally) relate to heuristic decision behavior. Our paper explores the importance of incentive contexts in shaping attention and the frequency of heuristic cognition.

There is research documenting that losses have larger effects than gains on physiological arousal. For example, it has been shown that subjects show higher arousal following losses using measures of skin conductance, heart rate, and pupil diameter (Hochman et al., 2009; Hochman and Yechiam, 2011; Yechiam and Hochman, 2013). There is also some neuropsychological evidence of dual-process reasoning in fMRI studies (Goel et al., 2000) and a large literature that examines the neural processing of gains and losses. Gehring and Willoughby (2002) for instance, find that the medial frontal cortex shows greater activation to losses, relative to gains, within 265 milliseconds of the reward stimuli. Sallet et al. (2007) and Foti et al. (2015) find evidence of distinct facets of reward processing related to gains and losses in the anterior cingulate cortex (ACC) and Fujiwara et al. (2009) find evidence of separate coding of monetary rewards and punishment in distinct subregions in the cingulate cortex. Similarly, Tom et al. (2007) find evidence that loss aversion is related to the same brain region responding differentially to gains and losses. Further, it is also thought that the ACC moderates the voluntary selection of behavior and encodes signals related to decision making (e.g., Amiez et al., 2005; Wallis and Kennerley, 2011). These studies are relevant as they highlight the differential neural responses to losses in regions of the brain that also influence decision making and attention allocation.

The main objective of our paper is to examine how the framing of incentives influences heuristic (S1) versus deliberative (S2) modes of thinking, which has been relatively unexplored. To do this we design three experimental treatments. In the “None” treatment, subjects receive no additional compensation for correct CRT answers, in the “Gain” treatment subjects are awarded an additional sum for each

<sup>2</sup>Considering other factors that might activate S2, Alter et al. (2007) examined whether the ease with which information comes to mind when facing a problem (i.e., the fluency of information) and the perceived difficulty of the judgement regulate when S2 will switch on. Similarly, Bourgeois-Gironde and Van Der Henst (2009) explore whether giving the decision-maker hints in the framing of the problem cues S2.

<sup>3</sup>Some recent work in area includes Oprea (2020), Bronchetti et al. (2023), and Oprea (2022).

correct response, and in the “Loss” treatment subjects who begin with an endowment are financially penalized for each incorrect response. In addition to other robustness checks, we conducted an auxiliary experiment to test the extent to which any treatment effects are on the extensive margin of triggering S2 and not just on the intensive margin of eliciting more cognitive effort. This second experiment was identical to the one just described except that it imposed a 20 second time constraint per question. This time constraint limits individuals’ ability to access S2 (e.g., Arsalidou et al., 2013; Bago and De Neys, 2017; Deck et al., 2021) and introduces a “cognitive load” that engages the working memory of our participants (Evans and Curtis-Holmes, 2005; Gillard et al., 2009; Kalyuga, 2011). We hypothesize that compromising S2’s ability to respond in this second experiment would attenuate any incentive treatment effects if they differ primarily in their ability to trigger S2.

## 2. Experiment 1

### 2.1. Methods

We recruited participants via Cloud Research and the experiment was conducted on Qualtrics in the summer of 2021.<sup>4</sup> To begin, subjects gave consent (the protocol was approved by the Middlebury College IRB), responded to basic demographic questions and were randomized into the different treatments. The treatment cells involved a different incentive prompt and a simple mathematical question to ensure subjects comprehended the incentives. Subjects who failed this comprehension question (9 of 610) were removed from the experiment.<sup>5</sup> In all incentive treatments subjects were awarded \$1.00 for completing the experiment. In the None treatment subjects received no additional compensation for correct CRT answers. In the Gain treatment, subjects were awarded an additional \$0.25 for each correct CRT response. And finally, in the Loss treatment subjects were endowed with \$2.75 and were penalized \$0.25 for each incorrect CRT response. To avoid any wealth effects, subjects were only informed about their performance after the completion of the experiment.<sup>6</sup>

After finishing the seven CRT questions, subjects responded to the NASA TLX survey, which was intended to measure work load. The survey instrument, developed by NASA’s Human Performance Group, asked for subjective rankings of the effort put into a task (on a 0–10 scale) across six demand dimensions (mental, physical, temporal, performance, effort, and frustration level). Because our task was entirely mental, we pruned the physical demand question but added another on attention demand.<sup>7</sup> On average, subjects spent approximately 5 minutes on the experiment and earned almost \$2, which translates to an hourly wage of approximately \$24/hr.

As noted, the experiment involved seven CRT questions. Three were the original CRT questions formulated by Frederick (2005).<sup>8</sup> Because these CRT questions are so ubiquitous, researchers have become concerned about repeated exposure. Toplak et al. (2014) propose four new CRT questions, which we also utilize.<sup>9</sup> Given the online implementation of the experiment, an additional concern is the ease at which subjects could search for answers online. To diminish this concern, we slightly modified the numerical values and changed the language in the questions. See Appendix Table B.1 for the seven CRT questions and our slightly modified versions.

<sup>4</sup>Subjects were based in the United States and were required to have prior experience in at least 100 HITs and have a HIT approval rating of 95%. An initial captcha question was utilized to eliminate bots.

<sup>5</sup>The three incentive prompts and the comprehension questions are reproduced in Appendix A.

<sup>6</sup>Our study was pre-registered at the AEA RCT registry (ID:AEARCTR-0007839).

<sup>7</sup>These questions are also reproduced in the Appendix.

<sup>8</sup>These original questions are the first three in Appendix Table B.1.

<sup>9</sup>We find some evidence that subjects perform better, on average, on the original CRT questions relative to the four new questions. Across incentive treatments the probability of a correct response on the original questions is 64% versus 47% on the new questions. This could reflect some prior exposure to the original CRT questions, however, even with the original questions more than a third of subjects still respond incorrectly.

**Table 1.** Experiment 1 participant characteristics.

|                 | Total  | Gain   | No reward | Loss   |
|-----------------|--------|--------|-----------|--------|
| Age             | 38.993 | 39.161 | 40.039    | 37.789 |
| Female          | 0.414  | 0.414  | 0.446     | 0.382  |
| Black           | 0.086  | 0.098  | 0.088     | 0.073  |
| Asian           | 0.085  | 0.072  | 0.054     | 0.127  |
| Hispanic        | 0.083  | 0.104  | 0.078     | 0.069  |
| Advanced degree | 0.113  | 0.132  | 0.138     | 0.103  |
| High income     | 0.556  | 0.570  | 0.529     | 0.556  |
| <i>N</i>        | 601    | 193    | 204       | 204    |

Note: Wald test for joint significance of characteristics: 15.35 ( $p = 0.35$ ).

## 2.2. Results

A total of 601 people participated in Experiment 1, more or less evenly split between the three incentive treatments (193 received gains, 204 received losses and 204 were not incentivized). Table 1 shows that the average participant was 39 years old, 41% were female, 9% were Black, 8% were Asian, 8% identified as Hispanic, 11% held advanced degrees (masters or above), and 56% reported having an income above \$50k per year. Considering treatment balance on these observables, a joint Wald statistic for all the characteristics predicting assignment to treatment is 15.35, suggesting we achieved randomization (regardless, we will present regression results with and without these controls).

Considering the choices that our participants made, the average person spent 311 seconds completing the experiment. On the CRT task, the overall average number of correct responses was 3.81 out of 7.<sup>10</sup> Finally, people earned an average of \$1.57 for their participation.

We begin our analysis by examining the extent to which our incentive treatments affected the amount of effort provided by our participants, measuring effort initially by the time spent on an individual CRT question and then by whether the individual correctly answered the question.<sup>11</sup> On the left side of Figure 1 in panel (a), we report the differences in the mean time spent on each CRT question. Overall, we see that incentives matter—participants spent more time when they could earn gains or take losses. Combining treatments, the average time spent increases a modest amount—4 seconds per question—when an incentive is provided ( $d = 0.15$ ,  $t(4205) = 3.40$ ,  $p < 0.01$ ), indicating participants are thinking longer about the questions when there are financial incentives.<sup>12</sup> Splitting the incentive treatments, we see that losses differentially affected how long a participant thought about the questions. While the Gain treatment increased the average time spent by 2.37 seconds per question compared to the control ( $d = 0.09$ ,  $t(2777) = 2.03$ ,  $p = 0.04$ ), the Loss treatment increased the average by 5.42 seconds per question ( $d = 0.20$ ,  $t(2854) = 4.00$ ,  $p < 0.01$ ) and the effect of losses is larger than the effect of gains ( $d = 0.09$ ,  $t(2777) = 2.05$ ,  $p = 0.04$ ).

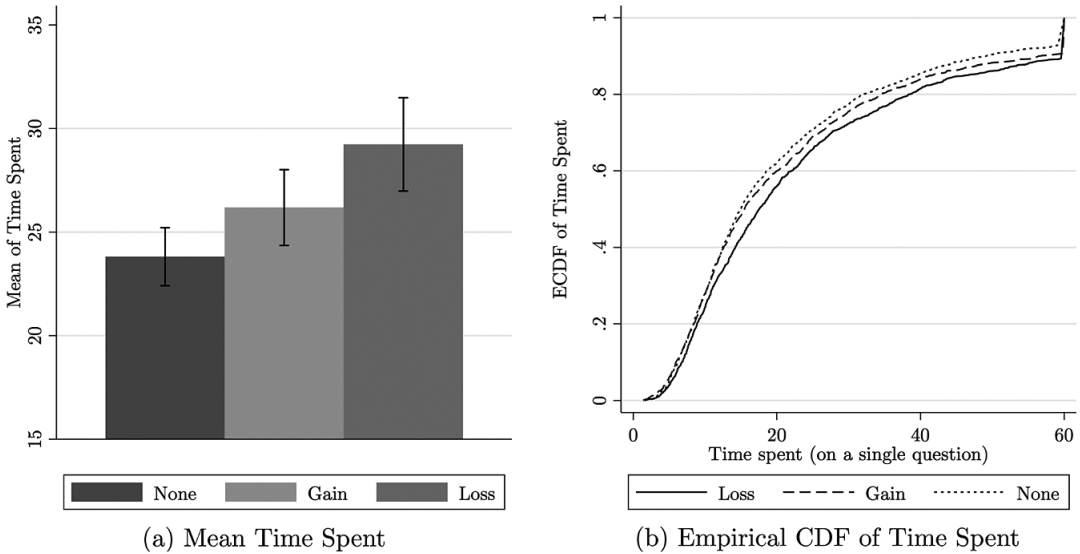
On the right of Figure 1, we examine the empirical cumulative distribution functions (ECDF) for the three incentive treatments (top coding at the 95 percentile or 60 seconds) and find that, because the ECDFs do not cross, the distributions have similar shapes (i.e., standard deviations) but different means. In fact, it is clear that the Loss treatment ECDF stochastically dominates both of the other treatments. These conclusions are supported by Kolmogorov–Smirnov tests ( $p < 0.01$  for both).

To more conservatively estimate the treatment effects of our incentive treatments, we regress the time spent answering a question on incentive treatment indicators in Table 2 (having no reward is the

<sup>10</sup>This rate of correct responses of 0.54 is somewhat larger than that reported in Toplak et al. (2014). They report a probability of correct response of 17 and 24% for the original 3 CRT questions and the four new CRT questions they propose, respectively.

<sup>11</sup>Various studies link the time spent making a choice to cognitive effort (e.g., Stanovich and West, 1998, 2000).

<sup>12</sup>When reporting t-test results, we include Cohen's  $d$  as a measure of effect size.



**Figure 1.** Time spent answering single CRT questions for the three incentive treatments (with 95% CI).

omitted category). In the first column, we replicate what Figure 1 and the associated *t*-tests indicated—implementing gains increased the average time spent on a question by 2.37 seconds and implementing losses increased this time by 5.42 seconds. Here we also see that the difference in these point estimates is 3 seconds ( $p = 0.04$ ). In the second column, we add controls for race, income, and education and find only small differences in our estimates as anticipated because our survey software correctly randomized participants to treatment. Finally, in the third column of Table 2, we account for the fact that each participant answered 7 of these questions by clustering our standard errors.<sup>13</sup> Accounting for this does increase the standard errors of our estimates somewhat and the difference between Gain and Loss becomes somewhat weaker ( $p = 0.14$ ), but our main finding, that losses increase the amount of time spent answering questions (i.e., effort), survives. In sum, like the previous literature, losses induce decision-makers to devote more effort to the task.

We find that losses also elicit more effort in terms of participants being more likely to answer the CRT questions correctly. In Figure 2, we see that our participants were slightly better than equally likely to answer the average CRT question correctly. In fact, the likelihood of getting a question correct in the None treatment is 0.498 and adding any incentive increases this by 7 percentage points on average ( $d = 0.14$ ,  $t(4205) = 4.32$ ,  $p < 0.01$ ). As with the time spent answering a question, when we separate the Gain and Loss treatments, we find that the participants who could suffer losses perform better. Here the effect of the Gain treatment is to increase the chance of answering correctly by 4.5pp ( $d = 0.09$ ,  $t(2777) = 2.40$ ,  $p = 0.02$ ), the effect of the Loss treatment is to increase this chance by 9.3pp ( $d = 0.19$ ,  $t(2854) = 5.02$ ,  $p < 0.01$ ) and the difference in these effects is 4.8pp ( $d = 0.10$ ,  $t(2777) = 2.54$ ,  $p = 0.01$ ).

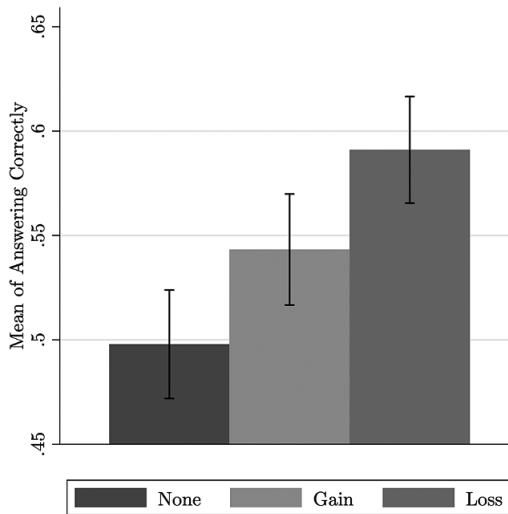
In Table 3, we estimate the treatment effects more carefully by regressing a correct answer indicator on the incentive conditions. Column (1) reproduces the results from Figure 2 and provides us an estimate for the treatment difference. Here we see that the Loss treatment effect is almost exactly twice as large as the Gain treatment effect ( $p < 0.01$ ). In the second column, we confirm that adding demographic controls changes these estimates only slightly, as expected, and in the third column we find that clustering the standard errors on the participant level, results in just the main effect of the Loss

<sup>13</sup>An alternative to clustering is to collapse data to the subject level, which we investigate below.

**Table 2.** The effect of gains and losses on time spent answering.

|                  | 1                            | 2                            | 3                            |
|------------------|------------------------------|------------------------------|------------------------------|
| Gain             | 2.372<br>(1.175)<br>[0.043]  | 2.271<br>(1.193)<br>[0.057]  | 2.271<br>(1.661)<br>[0.172]  |
| Loss             | 5.420<br>(1.354)<br>[0.001]  | 5.308<br>(1.364)<br>[0.001]  | 5.308<br>(1.772)<br>[0.003]  |
| Constant         | 23.813<br>(0.715)<br>[0.001] | 23.018<br>(0.938)<br>[0.001] | 23.018<br>(1.281)<br>[0.001] |
| Gain–loss        | –3.047<br>(1.480)<br>[0.040] | –3.037<br>(1.488)<br>[0.041] | –3.037<br>(2.034)<br>[0.136] |
| Controls         | No                           | Yes                          | Yes                          |
| Clustered errors | No                           | No                           | Yes                          |
| No. of obs.      | 4,207                        | 4,207                        | 4,207                        |

Note: Dependent variable is time spent answering (in seconds). Controls include race, income, and education. OLS with (robust standard errors) and [p-values] reported.



**Figure 2.** The likelihood of answering CRT questions correctly by incentive treatment (with 95% CI).

treatment persisting, as was the case with the time spent answering a question. Based on this measure of cognitive effort we also find that losses cause participants to work harder.

At this point, we have two pieces of evidence suggesting that losses differentially elicit more effort from decision-makers: participants who face losses think longer and are more likely to arrive at correct solutions to the CRT questions. While these results are consistent with the previous literature discussed above on losses and decision effort, we are mostly interested in whether losses can induce the decision-maker to move from S1 to S2 modes of thinking. Anticipating the difficulty of separating explanations based on cognitive effort from those based on cognitive process, we chose the CRT task because it also

**Table 3.** *The effect of gains and losses on thinking correctly.*

|                  | 1                            | 2                            | 3                            |
|------------------|------------------------------|------------------------------|------------------------------|
| Gain             | 0.045<br>(0.019)<br>[0.017]  | 0.049<br>(0.019)<br>[0.010]  | 0.049<br>(0.031)<br>[0.118]  |
| Loss             | 0.093<br>(0.019)<br>[0.001]  | 0.088<br>(0.019)<br>[0.001]  | 0.088<br>(0.030)<br>[0.004]  |
| Constant         | 0.498<br>(0.013)<br>[0.001]  | 0.495<br>(0.016)<br>[0.001]  | 0.495<br>(0.026)<br>[0.001]  |
| Gain–loss        | –0.048<br>(0.019)<br>[0.011] | –0.039<br>(0.019)<br>[0.040] | –0.039<br>(0.030)<br>[0.207] |
| Controls         | No                           | Yes                          | Yes                          |
| Clustered errors | No                           | No                           | Yes                          |
| No. of obs.      | 4,207                        | 4,207                        | 4,207                        |

Dependent variable is a correct response indicator.  
 Controls include race, income, and education. OLS with  
 (robust standard errors) and [*p*-values] reported.

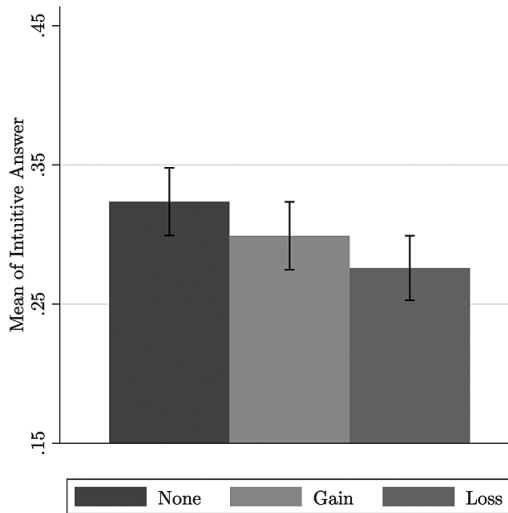
allows us to examine the frequency with which the incentive treatments elicit intuitive answers. Recall that there are three types of answers that a participant can give in the CRT: the correct answer, which may require the intervention of S2, and two types of incorrect answers, those that are intuitive and those that are simply wrong. Giving an intuitive (but incorrect) answer indicates a subject operating in S1.

In Figure 3, we illustrate how the incentive treatments affect the likelihood of responding to a CRT question with the intuitive, but incorrect, answer. In this case, adding a monetary incentive reduces the probability of responding intuitively by 3.6pp, overall ( $d = 0.08$ ,  $t(4205) = 2.44$ ,  $p = 0.02$ ). As in both of the previous cases, the Loss treatment is more effective at reducing intuitive responses—it more effectively triggers S2. The difference between the None and Gain treatments is just 2.5pp ( $d = 0.05$ ,  $t(2777) = 1.39$ ,  $p = 0.16$ ) but the difference between the None and Loss treatments is twice this magnitude, 4.8pp ( $d = 0.11$ ,  $t(2854) = 2.78$ ,  $p < 0.01$ ).

In Table 4, we first confirm the results of our summary tests and see that the differential efficacy of losses to trigger S2 thinking is robust to the addition of controls in column (2) and clustering the standard errors in column (3). In sum, our main results indicate that losses not only elicit more cognitive effort from decision-makers, they also cause them to move from fast and intuitive S1 thinking to slower, more deliberative, S2 thinking. Further, using similarly sized gains motivates participants too; however, the gain effect is approximately half as big as it is for losses. As a result, while losses robustly improve decision making over no incentives, their benefit over gains is less robust in our experimental paradigm.

To this point, we have focussed on only two of the three possible types of responses our participants can give to the CRT questions: correct answers and intuitive, but incorrect, answers. Participants also responded with incorrect answers that were not intuitive. Pooling across incentive treatments, 16% of the answers received were incorrect and not intuitive. In Figure 4, we illustrate the effect of incentives on these answers. Not only do incentives reduce the frequency of intuitive but wrong answers, they also have some effect on the frequency of other wrong answers. Compared to no incentives, gains reduce incorrect answers by 2.1pp ( $d = 0.06$ ,  $t(2777) = 1.47$ ,  $p = 0.14$ ), losses reduce incorrect answers by 4.6pp ( $d = 0.14$ ,  $t(2854) = 3.36$ ,  $p < 0.01$ ) and the difference in the incentive effects is 2.5pp ( $d = 0.07$ ,





**Figure 3.** The likelihood of answering CRT questions intuitively by incentive treatment (with 95% CI).

**Table 4.** The effect of gains and losses on thinking intuitively.

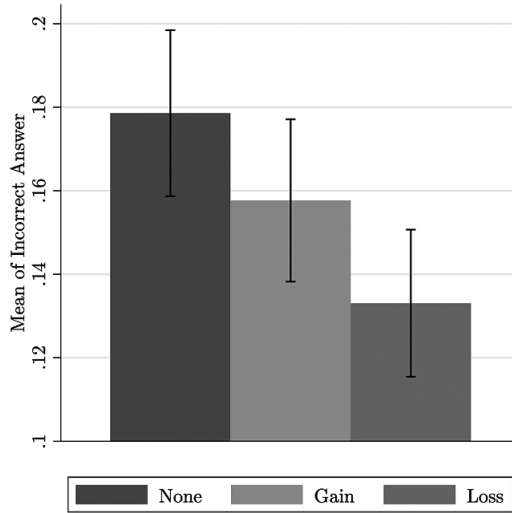
|                  | 1                            | 2                            | 3                            |
|------------------|------------------------------|------------------------------|------------------------------|
| Gain             | -0.024<br>(0.018)<br>[0.163] | -0.026<br>(0.018)<br>[0.145] | -0.026<br>(0.026)<br>[0.324] |
| Loss             | -0.048<br>(0.017)<br>[0.005] | -0.045<br>(0.017)<br>[0.009] | -0.045<br>(0.025)<br>[0.077] |
| Constant         | 0.324<br>(0.012)<br>[0.001]  | 0.319<br>(0.015)<br>[0.001]  | 0.319<br>(0.021)<br>[0.001]  |
| Gain–loss        | 0.023<br>(0.017)<br>[0.178]  | 0.019<br>(0.017)<br>[0.264]  | 0.019<br>(0.026)<br>[0.458]  |
| Controls         | No                           | Yes                          | Yes                          |
| Clustered errors | No                           | No                           | Yes                          |
| No. of obs.      | 4,207                        | 4,207                        | 4,207                        |

*Note:* Dependent variable is an intuitive response indicator. Controls include race, income, and education. OLS with (robust standard errors) and [*p*-values] reported.

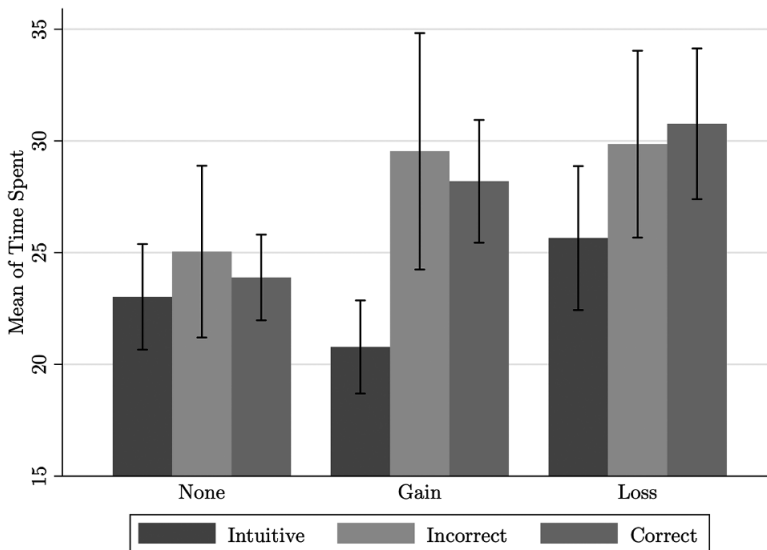
$t(2777) = 1.84, p = 0.06$ .<sup>14</sup> We now consider results that include all three types of answers to get a better sense of how the incentives might have affected the transitions made by the participants from one type of answer to another.

As participants devote more cognitive effort to the CRT questions, there are two paths that they can take on their way to the correct answer. First, using little effort, people can simply be wrong. They can then stumble across the intuitive answer after putting in a bit more effort and arrive at the correct

<sup>14</sup>It is worth mentioning that if we pool the data from this experiment with the one discussed in section 3, all these differences are considerably stronger ( $p < 0.01$  for all).



**Figure 4.** Incentives and the frequency of unintuitive incorrect answers (with 95% CI).



**Figure 5.** Time spent by type of answer and incentive.

answer once they have devoted enough effort to the task. Alternatively, decision-makers who think very little about the question might quickly land on the intuitive answer but find themselves at an unintuitive answer that is still wrong once they put in more effort. These people might sense that the intuitive answer seems too simple and devote more resources to the problem, eventually arriving at the correct answer after struggling a bit with the wrong answer. In the first explanation, additional cognitive effort leads them through: *incorrect* → *intuitive* → *correct*, while the second explanation swaps the first two states: *intuitive* → *incorrect* → *correct*.

Our proxy for cognitive effort is time spent and in Figure 5 we report the mean time spent that led to each of the answer types across the three incentive conditions. Pooling across incentives, participants spent the least time to arrive at the intuitive answer (23.12 seconds) and they spent about the same amount of time on incorrect and correct answers (27.89 and 27.80 second, respectively). Though not

**Table 5.** Effort differences by answer type and incentive.

|             | 1                            | 2                            | 3                            |
|-------------|------------------------------|------------------------------|------------------------------|
| Incorrect   | 1.982<br>(2.303)<br>[0.390]  | 8.745<br>(2.817)<br>[0.002]  | 4.126<br>(2.693)<br>[0.126]  |
| Correct     | 1.023<br>(1.546)<br>[0.508]  | 6.784<br>(1.782)<br>[0.001]  | 5.398<br>(2.428)<br>[0.026]  |
| Constant    | 23.174<br>(1.388)<br>[0.001] | 17.406<br>(1.648)<br>[0.001] | 25.682<br>(2.034)<br>[0.001] |
| Controls    | Yes                          | Yes                          | Yes                          |
| Gains       | No                           | Yes                          | No                           |
| Losses      | No                           | No                           | Yes                          |
| No. of obs. | 1,428                        | 1,351                        | 1,428                        |

*Note:* Dependent variable is time spent answering (in seconds). Controls include race, income, and education. OLS with (robust standard errors) and [*p*-values] reported.

definitive, the fact that intuitive answers were given considerably quicker than the other two types is evidence in favor of the second explanation described above ( $d = 0.18$ ,  $t(1916) = 3.37$ ,  $p < 0.01$  and  $d = 0.17$ ,  $t(3547) = 3.72$ ,  $p < 0.01$ ). At first blush, the fact that the differences in the time spent across answer types when there are no incentives are not significant seems to be contrary to a model in which the transition to S2 (and more effort) is necessary to get to the correct answer. However, this interpretation is confounded by the fact that there are considerably more correct answers with incentives. Put differently, in a population that is heterogeneous in the processing time needed to get to the correct answer, the clever people will do so as quickly as people who offer the intuitive answer (and would need more time to get it right). This will be true with financial incentives too, but the difference is that the incentives entice people who are not quite as clever to spend more time and some of them will get it right after more deliberative thinking in S2. Further, Figure 5 suggests that incentives act to increase the time spent by our participants as they reason to both incorrect and correct answers, but not intuitive ones. These results are confirmed in Table 5 and are also consistent with transitioning from quick intuitive answers that take little effort to conjure to incorrect and correct answer after further contemplation. As the point estimates elucidate, our participants spend about the same amount of time generating all three types of answers when there are no incentives. Gains elicit more effort that results in incorrect and correct answers but they do not increase effort to come up with the intuitive answer. Losses have similar effects on incorrect and correct answers as gains, but they also slightly increase the amount of time spent on intuitive answers.

### 2.3. Robustness

To explore the robustness of our main results, we first aggregate the data for each individual and examine the total time spent solving CRT questions, the total number of questions our participants answer correctly and the total number of intuitive responses our participants give. We then look for heterogeneous treatment effects of our incentives among a subgroup of our participants known for a predisposition toward S1 reasoning young men.

An analysis of the aggregated data can be found in Table 6. In the first two columns we examine the total time spent answering CRT questions, first without controls and then with them. In both cases,

**Table 6.** Treatment effects using aggregated data.

|             | 1                              | 2                              | 3                            | 4                            | 5                            | 6                            |
|-------------|--------------------------------|--------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|
| Gain        | 16.607<br>(11.523)<br>[0.150]  | 15.898<br>(11.694)<br>[0.175]  | 0.087<br>(0.061)<br>[0.151]  | 0.096<br>(0.060)<br>[0.113]  | -0.079<br>(0.084)<br>[0.349] | -0.078<br>(0.084)<br>[0.352] |
| Loss        | 37.938<br>(12.283)<br>[0.002]  | 37.159<br>(12.478)<br>[0.003]  | 0.171<br>(0.057)<br>[0.003]  | 0.163<br>(0.057)<br>[0.004]  | -0.159<br>(0.086)<br>[0.064] | -0.151<br>(0.085)<br>[0.077] |
| Constant    | 166.691<br>(6.390)<br>[0.001]  | 161.124<br>(9.021)<br>[0.001]  | 1.249<br>(0.045)<br>[0.001]  | 1.243<br>(0.051)<br>[0.001]  | 0.817<br>(0.056)<br>[0.001]  | 0.803<br>(0.069)<br>[0.001]  |
| Gain-loss   | -21.331<br>(14.213)<br>[0.134] | -21.261<br>(14.323)<br>[0.138] | -0.084<br>(0.054)<br>[0.122] | -0.067<br>(0.054)<br>[0.217] | 0.080<br>(0.090)<br>[0.372]  | 0.073<br>(0.090)<br>[0.416]  |
| Dep. var.   | Time                           | Time                           | Correct                      | Correct                      | Intuitive                    | Intuitive                    |
| Controls    | No                             | Yes                            | No                           | Yes                          | No                           | Yes                          |
| No. of obs. | 601                            | 601                            | 601                          | 601                          | 601                          | 601                          |

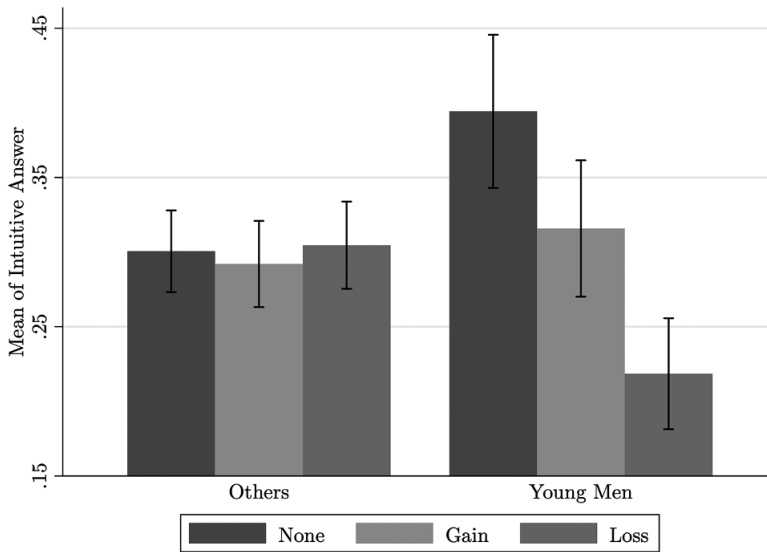
*Note:* Dependent variable varies between total time spent, number of correct responses and number of intuitive responses. Controls include race, income, and education. OLS or negative binomial with (robust standard errors) and [*p*-values] reported.

we see that aggregating the data has two effects. First, the point estimate on the Gain treatment—spending 16.61 seconds more than in the None treatment—is sizable, but the standard error rises such that  $p = 0.15$ . Second the point estimate on the Loss treatment—an additional 37.94 seconds above what is spent in the control—remains large, even when the covariates are added. Columns (3) and (4) report the results of analyzing the number of correct responses to the CRT questions. Again, we find in these two columns that regardless of whether controls are added to the regression or not, the standard error on the effect of the Gain treatment rises while this doesn't happen with the sizable Loss treatment effect. In the final two columns of Table 6, we see the same pattern as in the previous columns. When considering the number of intuitive responses, the Gain treatment effect loses predictive power in the aggregated data but the Loss effect remains strong. The punchline of this analysis is that only the Loss treatment effects are robust to aggregating the data across individual participants. Most importantly, we continue to find evidence that losses trigger S2 thinking.

Because the benefit of financial incentives may be more (or less) necessary with different subgroups within the population, we also split the sample in a way consistent with the literature on brain development to test whether we are more likely to find effects where we would expect. Specifically, mounting evidence suggests that brains develop slowly and differentially by sex (Cowell et al., 2007). It is now thought that young men, in particular, have trouble with planning and other executive functions associated with working memory and the prefrontal cortex (Kaller et al., 2012; Paus, 2005). Based on these findings, we hypothesized that if losses truly help decision-makers switch into S2, their effects should be more pronounced among people who might find this transition difficult.<sup>15</sup> That is, if older people, for example, are predisposed to a higher likelihood of S2 thinking, or who switch to S2 more readily, they may not need the extra stimulus provided by losses (required by younger people), to access S2.<sup>16</sup> Young men, by contrast, are expected to give more intuitive (S1) answers in the baseline without incentives and therefore have more latitude to switch when presented with the loss stimulus.

<sup>15</sup>Note that this aspect of our results was not included in our pre-analysis plan and is considered exploratory.

<sup>16</sup>Indeed, our data suggest that older (or female) participants do spend more time on the questions, especially when there are no incentives ( $d = 0.17$ ,  $t(1426) = 2.33$ ,  $p = 0.02$ ). Additionally, these other participants are considerably more likely to respond



**Figure 6.** The likelihood of answering CRT questions intuitively separately for young men (with 95% CI).

To assess whether the effects of losses on switching to S2 are more pronounced among young men, we split the sample into two groups, men under the age of 35 (the median age of our male participants) and the rest, resulting in the classification of 27% of the sample as young men. In Figure 6, we see that the treatment effects we found in the previous section are muted among the “Others.” For the other participants on the left of the figure, gains reduce the number of intuitive responses by just 0.85pp and losses actually increase them slightly by 0.41pp. The results differ for young men. Here, on the right of Figure 6, we see that young men respond more intuitively than the others, overall. Considering the treatment differences for young men, we find that gains reduce the frequency of intuitive responses by 7.80pp ( $d = 0.17, t(747) = 2.25, p = 0.02$ ) and losses reduce intuitive responses by a considerable 18pp ( $d = 0.43, t(824) = 5.58, p < 0.01$ ). Put differently, young men do offer more intuitive answers, as expected, and the treatment effect of losses is much greater among young men than it is in the rest of the population.<sup>17</sup>

Combining all the data in a regression setting with interactions, in Table 7, we confirm in column (1) that among the other participants, the treatment effects largely vanish. Both point estimates are less than 1pp. Further, these point estimates for the other participants change little when we include our controls for race, income, and education. The interaction of being young, male and being in the Loss treatment is large however, regardless of what controls are added. Near the bottom of Table 7, we report the estimated treatment effects for the young men. For the Gain treatment the estimated effects for young men are between 6.2pp and 7.8pp and they grow to between 16.7pp to 17.6pp in the Loss treatment.<sup>18</sup>

with the correct answer when there are no incentives ( $d = 0.18, t(1426) = 2.87, p < 0.01$ ). Both of these results suggest that these “other” participants spend more time in S2 or require less of a financial nudge to transition into S2.

<sup>17</sup>As we decrease the cutoff age toward 25, the age at which the brain is thought to be mostly developed, the results in Figure 6 persist even as the target sample shrinks.

<sup>18</sup>It is important to note that our results along gender dimensions, in general, are also consistent with existing literature. Two meta-analyses on CRT performance find that men, overall, tend to perform better than women (Brañas-Garza et al., 2019; Yechiam and Zeif, 2023). Examining this in our data, we find that subjects who self-report as female get 0.78 fewer CRT questions correct ( $p < 0.01$ ).

**Table 7.** *The differential effect of losses on thinking intuitively among young men.*

|                           | 1                            | 2                            |
|---------------------------|------------------------------|------------------------------|
| Gain                      | −0.009<br>(0.020)<br>[0.674] | −0.017<br>(0.021)<br>[0.422] |
| Loss                      | 0.004<br>(0.020)<br>[0.842]  | 0.005<br>(0.021)<br>[0.803]  |
| Young man                 | 0.094<br>(0.030)<br>[0.002]  | 0.082<br>(0.030)<br>[0.007]  |
| Gain × young man          | −0.070<br>(0.040)<br>[0.084] | −0.046<br>(0.041)<br>[0.267] |
| Loss × young man          | −0.180<br>(0.038)<br>[0.001] | −0.173<br>(0.039)<br>[0.001] |
| Constant                  | 0.301<br>(0.014)<br>[0.001]  | 0.300<br>(0.016)<br>[0.001]  |
| Gain + (gain × young man) | −0.078<br>(0.039)<br>[0.025] | −0.062<br>(0.053)<br>[0.079] |
| Loss + (loss × young man) | −0.176<br>(0.035)<br>[0.001] | −0.167<br>(0.033)<br>[0.001] |
| Controls                  | No                           | Yes                          |
| No. of obs.               | 4,207                        | 4,207                        |

*Note:* Dependent variable is an intuitive response indicator.  
 Controls include race, income, and education.  
 OLS with (robust standard errors) and [*p*-values] reported.

### 3. Experiment 2

To sharpen our results, we ran a second version of the experiment and added a time constraint to each CRT question—a common experimental manipulation to reduce the impact of deliberation. In principle, if losses differentially trigger S2 thinking, when we impose a time constraint on our participants (making it harder to switch to S2), the monetary incentive treatment effects should vanish.

Because S1 processes are thought to be automatic and intuitive, they tend to be fast, while S2 processes are deliberative, reflective, and slow. Time constraints, therefore, are intended to increase reliance on intuition by cutting reflective processes short. In fact, Frederick (2005) reports that correct responders in the CRT often considered the incorrect but intuitive answer first. This is consistent with our results from Experiment 1 (e.g., Figure 5) and Travers et al. (2016), who used mouse tracking to study the time course of reasoning on the CRT questions and found that participants were initially drawn to the intuitive but incorrect answers, even when the correct answers were ultimately chosen. Further, Finucane et al. (2000), who were among the first to use time constraints to limit S2 deliberation, show that time constrained participants were more likely to use the affect heuristic to judge the risk and benefits of specific hazards. Similarly, Evans and Curtis-Holmes (2005) imposed a 10-second

time constraint on answering syllogisms and found that the constrained group had more difficulty discriminating between valid and invalid arguments, evidence of belief bias, and there not being enough time to access S2. Subsequent studies have observed similar effects of time constraints on the incidence of other intuitive thinking like the natural number bias (i.e., thinking 7/11 is larger than 4/5 because 7 is larger than 4 and 11 is larger than 5) (Van Hoof et al., 2020), a pattern confirmed in a recent comprehensive experiment by Isler and Yilmaz (2023), which finds that time constraints are the most effective way to activate intuition (and hinder reflection).<sup>19</sup>

### 3.1. Methods

The only difference between Experiment 1 and 2 is that we restricted the time allocated to each question to 20 seconds. The choice of 20 seconds was a tradeoff between giving subjects enough time to read and comprehend the questions, but also making them feel rushed. The average reading speed of an adult is around 250 words per minute, which would result in a reading time of 13.2 seconds for our longest question. The median question response time in our unconstrained experiments was 16 seconds. So the 20 second limit (indicated by a clock counting down) seemed to be a reasonable choice to allow question comprehension, but also cause subjects to feel pressured. Manipulating response time is a standard method for impeding S2. In the context of answering questions like the CRT, Gillard et al. (2009) showed that reducing the time spent on similar questions to the median time spent by unconstrained participants considerably reduced correct answers (and increased intuitive responses). More broadly, Kalyuga (2011) reviews the literature on working memory and explains how time constraints are an essential part of any cognitive load.

During the summer of 2021, we recruited an additional 609 participants from Cloud Research for Experiment 2 (on this round 10 failed the comprehension quiz). As before, subjects gave consent, responded to demographic questions, were randomized to an incentive treatment, attempted the CRT questions, and responded to the NASA TLX survey. Like the participants in Experiment 1, these additional participants spent about 5 minutes on the experiment and earned around \$2, on average.

### 3.2. Results

As seen in Table 8, the participants of Experiment 2 were (roughly) evenly divided into the three conditions: 199 were sorted into the Gain treatment, 202 faced losses and 208 were given no additional reward for their correct CRT answers. Comparing Tables 1 and 8, we see that the demographics of the second experiment match those of the first. Considering treatment balance on these observables, a joint Wald statistic for all the characteristics predicting assignment to treatment is 13.83 ( $p = 0.46$ ), consistent with achieving randomization to treatment.

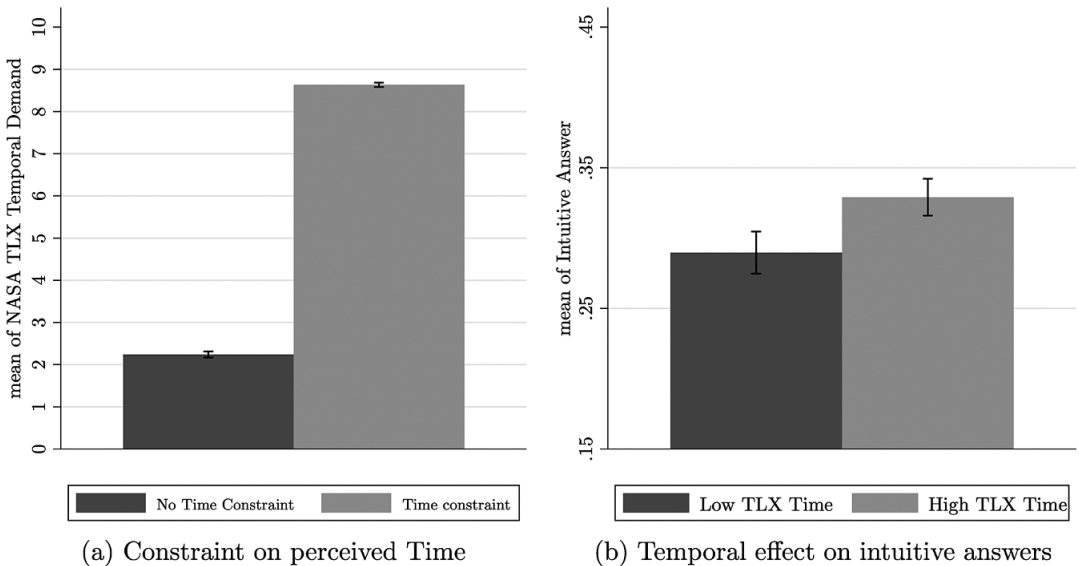
Because we expected that implementing a time constraint will hinder subjects' access to S2, regardless of the incentive treatment, we anticipated that the Loss treatment would lose its ability to trigger S2 in this second experiment where it is, more or less, unavailable. We begin our analysis of the data from this second experiment by testing whether the time constraint manipulation had the intended effect. On the left of Figure 7, we see that imposing the time constraint dramatically increased the reported feeling of hurry or being rushed (measured by NASA's TLX temporal demand item). The average time pressure experienced by participants in the no time constraints experiment is relatively mild—2.24 out of a possible 10—but the load reported by the participants that answered the CRT questions with time constraints is considerably more severe—8.63 out of 10 ( $d = 3.66$ ,  $t(1208) = 53.41$ ,  $p < 0.01$ ). This difference suggests that the constraint did affect participants' impression of how much

<sup>19</sup>An alternative interpretation is that time constraints could nudge participants to focus, activating S2. In this case the constraint should lead to fewer intuitive answers and possibly more correct ones because participants transition to S2 more quickly. Our data favor the standard interpretation, however. We find that intuitive answers come quickest (Figure 5) and the time constraint leads to more intuitive, and fewer correct, responses (Figures 7 and 8 below). These differences in the frequency of intuitive and correct answers between the no time constraint and time constraint experiments are significant at the 1% level.

**Table 8.** Experiment 2 participant characteristics.

|                 | Total  | Gain   | No reward | Loss   |
|-----------------|--------|--------|-----------|--------|
| Age             | 39.112 | 39.497 | 40.346    | 37.460 |
| Female          | 0.470  | 0.452  | 0.490     | 0.465  |
| Black           | 0.082  | 0.075  | 0.063     | 0.109  |
| Asian           | 0.062  | 0.060  | 0.063     | 0.064  |
| Hispanic        | 0.077  | 0.101  | 0.072     | 0.059  |
| Advanced degree | 0.161  | 0.144  | 0.138     | 0.124  |
| High income     | 0.540  | 0.527  | 0.577     | 0.515  |
| <i>N</i>        | 609    | 199    | 208       | 202    |

Note: Wald test for joint significance of characteristics: 13.83 ( $p = 0.46$ ).



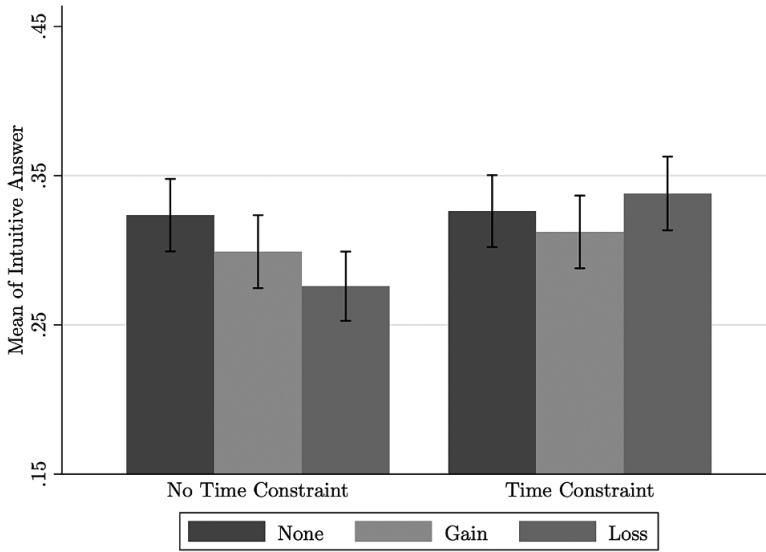
**Figure 7.** Manipulation check: does the time constraint affect S2 access?

time pressure they felt and, presumably, how much they could consult S2. To corroborate this inference, in the right panel of Figure 7, we see that participants who report higher than median NASA TLX temporal scores were 4pp more likely to respond to CRT questions with the intuitive response ( $d = 0.09$ ,  $t(8468) = 3.85$ ,  $p < 0.01$ ).<sup>20</sup>

Comparing between time constraint treatments, we see the anticipated results in Figure 8. On the left, we reproduce our main results from Experiment 1, without any time constraint (i.e., Figure 3), in which both gains and losses affect the frequency of intuitive CRT responses but losses reduce the frequency twice as much. On the right of Figure 8, we illustrate the results of the second experiment with a constraint. Averaging across the incentives, we see that the time constraint did result in more intuitive answers overall relative to Experiment 1 with no time constraint ( $d = 0.06$ ,  $t(8468) = 2.59$ ,  $p < 0.01$ ). Most importantly, we also find that the treatment effects fade with the imposition of a time constraint: the Gain treatment reduces intuitive responses by just 1.4pp ( $d = 0.03$ ,  $t(2847) = 0.80$ ,

<sup>20</sup>We get similar results when we focus on the TLX item that is intended to measure anxiety or stress or if we simply aggregate across all the TLX responses.





**Figure 8.** A time constraint hindering S2 attenuates the treatment effects.

$p = 0.42$ ) and the Loss treatment actually increases intuitive responses slightly by 1.2pp ( $d = 0.03$ ,  $t(2868) = 0.67$ ,  $p = 0.50$ ).

In Table 9, we combine the data and estimate treatment effects for both experiments (i.e., no time constraint and time constraint). As in Table 4, we see that the Loss treatment reduces the chance that a participant will respond intuitively (4.8pp or 4.6pp depending on whether controls are added or not) but the Gain treatment does not affect intuitive responses nearly as much. Adding the point estimate of the time constraint interaction term at the bottom of Table VII, reveals that the loss treatment no longer triggers S2 when access to it is hindered. In other words, our results indicate that the prospect of losses does not just elicit more effort from decision-makers, it helps them transition to deliberative S2 thinking and this disappears when S2 is unavailable.

Given the heterogeneous treatment effects discussed above, the final aspect of robustness to check is to make sure that imposing the time constraint also enervates the incentive treatment effects among the participants most predisposed to makes choices from S1—the young men. In Figure 9, we see the same overall pattern for the other participants as we saw in Figure 8. Among the other participants working with the time constraints, the financial incentives do not have a large effect on replying with the intuitive response. For the young men on the right of Figure 9, we also see that the treatment effects diminish substantially when the constraint is imposed. Whereas gains reduced intuitive responses 7.8pp for young men who face no time constraint, the reduction is only 3.4pp for the constrained young men ( $d = 0.07$ ,  $t(670) = 0.90$ ,  $p = 0.37$ ). Considering the more substantial effect of losses, unconstrained young men responded with the intuitive response 18pp less but this effect shrinks by more than half to just 7pp when they are constrained ( $d = 0.15$ ,  $t(642) = 1.885$ ,  $p = 0.06$ ). Even among this subpopulation, we find that losses do not just elicit more effort, they differentially push people toward S2 thinking.

#### 4. Discussion

The dual system approach has been a prominent conceptual framework for understanding choice (Evans, 2008; Kahneman, 2011). However, little is known about the factors that cause individuals to shift between S1 and S2 (De Neys, 2023). Our experiment, designed to specifically examine S1 and S2

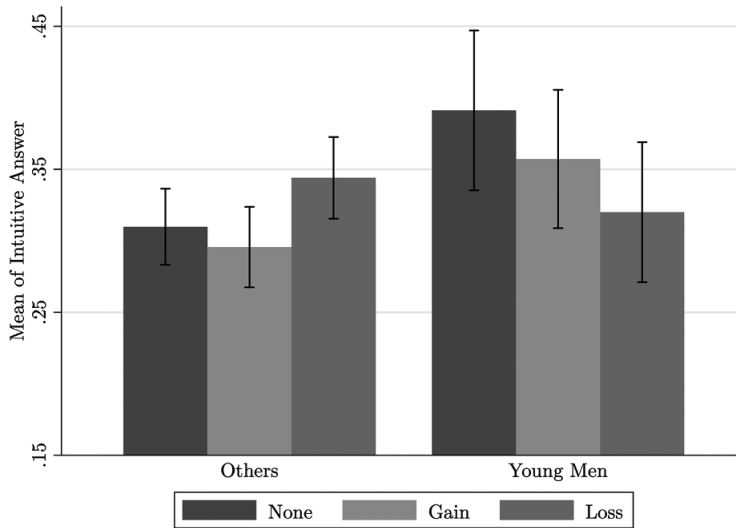
**Table 9.** *The differential effect of a time constraint on thinking intuitively.*

|                      | 1                            | 2                            |
|----------------------|------------------------------|------------------------------|
| Gain                 | −0.024<br>(0.018)<br>[0.163] | −0.025<br>(0.018)<br>[0.149] |
| Loss                 | −0.048<br>(0.017)<br>[0.005] | −0.046<br>(0.017)<br>[0.008] |
| Time constraint      | 0.003<br>(0.017)<br>[0.877]  | 0.004<br>(0.017)<br>[0.796]  |
| Gain × time          | 0.011<br>(0.025)<br>[0.671]  | 0.015<br>(0.025)<br>[0.554]  |
| Loss × time          | 0.059<br>(0.025)<br>[0.016]  | 0.054<br>(0.025)<br>[0.027]  |
| Constant             | 0.324<br>(0.012)<br>[0.001]  | 0.328<br>(0.014)<br>[0.001]  |
| Gain + (gain × time) | −0.014<br>(0.017)<br>[0.424] | −0.011<br>(0.017)<br>[0.538] |
| Loss + (loss × time) | 0.012<br>(0.017)<br>[0.502]  | 0.009<br>(0.017)<br>[0.616]  |
| Controls             | No                           | Yes                          |
| No. of obs.          | 8,470                        | 8,470                        |

*Note:* Dependent variable is an intuitive response indicator. Controls include race, income, and education. OLS with (robust standard errors) and [*p*-values] reported.

thinking, highlights that the framing of rewards partially determines whether subjects provide intuitive, but incorrect, responses (S1) or engage in more effortful, contemplative thinking (S2). We also find that these framing effects are important when subjects have a tendency to occupy S1 thinking—if subjects have a tendency to operate in S2 by default, the framing effects go away since there is no shift needed to invoke S2. These results are consistent with a view of switching that places emphasis on external cues (à la Evans and Stanovich, 2013), in our case losses versus gains and their implications for the expected magnitude of the payoffs at stake. In these models, decision-makers often default to the (potentially biased) S1 response because they simply fail to detect the need to activate S2 (Morewedge and Kahneman, 2010; Stanovich and West, 2000). Our results suggest that because similarly sized losses are often valued, and perhaps noticed, more readily than gains, the external framing of a choice (i.e., the choice architecture) may be manipulated to help decision-makers cross the stimulus/attention threshold necessary to invoke S2.

Our results also speak to the broader issues on the efficacy of behaviorally inspired public policy. Important policy institutions have urged decision makers to rely on S2 thinking, and this emphasis might shape policy design (e.g., use of policy “nudges”) in an effort to improve decisions (Melnikoff



**Figure 9.** Comparing the impact of incentives under a time constraint between young men and others.

and Bargh, 2018). However, in understanding the efficacy of these policy nudges, context clearly matters (Chater, 2018). A naive conclusion from our results would be that policymakers should frame decisions and incentives in terms of losses/costs and this would improve decision outcomes. However, this clearly depends on the degree to which individuals are approaching decisions in an S1 state of mind. If they are already operating in more of a S2 cognitive state, the losses nudge may be irrelevant in their decisions. Our heterogeneous results highlight this nicely. For individuals who have a higher propensity to operate in S1 (young men) the loss nudge is highly efficacious. However, for subpopulations with a higher propensity to operate in S2 the nudge is irrelevant. In the context of designing behavior public policy through an S1/S2 lens, it is clearly important to understand if the decision environment is one in which individuals would have a tendency to rely on S1, and our results highlight that this may be informed by understanding the characteristics of the decision makers.

Considering other models of decision making that focus more on the transition from incorrect answers to correct ones, we see our experiment as suggesting that a hybrid approach might be productive. These alternative models (e.g., Aha! Kounios and Beeman, 2009) have a hard time differentiating between the different ways that a decision-maker can get it wrong while this is a strength of the dual processing model. Another way of framing this is that both the Aha! model and dual process model emphasize binary outcomes: incorrect versus correct or intuitive versus correct, respectively. However, a more complete model of cognition would describe the assemblage of incorrect, intuitive (but incorrect), and correct responses.

While the effect sizes are modest, our data is consistent with the view that all three of these outcomes are relevant. Consider two stylized hybrid models of cognition hinted at above and what happens as the stakes of the decision are raised and cognitive effort is deployed. In the first, the decision-maker starts by providing little effort when the stakes are low and is likely to make an incorrect choice. However, as the stakes increase, they engage in more effort and are attracted to the intuitive answer, though it is also incorrect. However, if the stakes/stimulus are strong/high enough, the decision-maker puts in even greater effort and converges on the correct answer. We have Model 1: Low Incentives | Low Effort | Incorrect Choice → Medium Incentives | Medium Effort | Intuitive (but incorrect) Choice → High Incentives | High Effort | Correct Choice. An alternative model (2) that is somewhat more consistent with dual processing (and Aha!) switches the order of the first two states. Here when the stakes are perceived to be low, the decision-maker responds with the quick intuitive (but effortless) answer and sometimes this is wrong. However, when the stakes increase, the decision-maker transitions from S1

to S2 (as in the diffusion model of Alós-Ferrer, 2018) but this does not imply that she gets it right. While in S2, the decision-maker has to put in enough effort to potentially have the Aha! moment and transition from being incorrect to correct.

In this case, we have Model 2: Low Incentives | Low Effort | Intuitive (but incorrect) Choice → Medium Incentives | Medium Effort | Incorrect Choice → High Incentives | High Effort | Correct Choice. In the first model, incorrect responses could be thought of as stemming from noise or very limited cognitive inputs. In the second model, incorrect responses could be thought of as stemming from cognitive engagement but a failure to have an Aha! breakthrough.

Using decision time as a proxy for cognitive effort, we provide preliminary evidence as to which of these models is better supported by our data. Recall from Section 2.2, that our incentive manipulations increase effort, but the right two-thirds of Figure 5 are consistent with those efforts being directed disproportionately toward the production of incorrect and correct answers.

In other words, these results are more consistent with the second model: when incentives exist, subjects shift into higher effort S2, and once in S2 there is the possibility of an Aha! type moment where they arrive at the correct solution and effort stops. The statistically indistinguishable amount of time spent on incorrect and correct answers is suggestive of this type of effort halting process: once in S2, two decision makers may spend the same amount of additional time on a puzzle, but it is somewhat up to chance whether each gets it right. While we view these results as redolent, we emphasize that this evidence is certainly not conclusive regarding which type of cognitive model is the “correct” one. If nothing else, however, these results highlight that hybrid models with three response states (incorrect, intuitive, correct) may be important in developing a richer understanding of cognitive processes and decision making.

We also see our experiment and results as dovetailing with the recent influential literature on complexity and attention. As Oprea (2020) for instance shows, decision-makers avoid complexity precisely because cognitive effort is costly. As a result, they are more likely to make poor choices unless external cues are used to frame (and perhaps change the perceived magnitude of) the incentives so as to trigger deliberative reasoning. Similarly, Bronchetti et al. (2023) in the rational inattention literature, find evidence that decision-makers undervalue the benefits of paying close attention to choices—it’s not that they understand the implications of distractions and choose optimally, they objectively pay too little attention to the choices at hand. Because people do not optimally allocate their attention, there are surely benefits of designing interventions to help them get closer to what would be ideal. Again, our results indicate that, in the right context, simple framing might help “nudge” people across their attention threshold and help them accrue substantial welfare gains.

Since dual system thinking is often offered as an explanation for biased decision making (e.g., Kahneman, 2011), our results may help highlight reward mechanisms that prod individuals toward S2 thinking and, thus, better outcomes. Our results also align nicely with the neuro-imaging literatures that highlight the differential neural processing of gains and losses in regions of the brain that also moderate decision making and attention (e.g., Foti et al., 2015; Fujiwara et al., 2009). Interesting follow-up work would be to examine the neural activity related to S1 and S2 thinking, and our results highlight an experimental mechanism to differentially elicit these types of decision making.

Our results may also motivate more thinking on the mechanisms behind differential responses to gains and losses (i.e., loss aversion). Concerning managerial decision-making, to what extent are the differences observed in Hossain and List (2012), for instance, wherein conditional incentives framed as losses increased productivity on the factory floor more than equally sized gains, the result of workers spending more time in a deliberative mindset? More generally, Oprea (2022) shows that subjects display loss averse type behavior even when the choice task is manipulated to remove losses. The findings suggest that much of the “loss averse” behavior observed may be driven by the complexity of evaluating lotteries rather than risk preferences. These results highlight that “loss averse” type behavior can possibly be generated by how limited-attention subjects respond to complex decision tasks.

We find evidence that the framing of financial rewards triggers different types of attention and thinking, and thus responses. In the vast literature observing different responses to gambles spanning

gains and losses, an important question relates to the degree to which these responses stem from different modes of thinking about the task versus some subjective preferences over gains and losses. There is some important work investigating these types of questions. For example, Yechiam and Hochman (2013) highlight evidence that is consistent with an interpretation that the strong effect of losses on behavior is not necessarily a result of losses being given more weight in decisions, but rather stems from the fact that losses have distinct effects on performance via increased attention and arousal. They present a loss-attention model where losses impact behavior by increasing on-task attention, which increases the likelihood of responding in a manner that is less random and more consistent with the task reinforcement structure. Using data from process tracking, Lejarraga et al. (2019) find evidence that people invest more attentional resources when evaluating losses than when evaluating gains, even when their choices do not reflect loss aversion. Likewise, Gehring and Willoughby (2002) find increased activation in the medial frontal cortex to losses relative to gains, but they also do not find evidence of loss aversion in the average study participant. Given that loss aversion is such a ubiquitous theory, future research deepening our understanding of the relative importance of attention versus subjective preferences in driving differential behavior in loss domains is certainly warranted. Studies that examine loss aversion without controlling for differential levels of attention confound these two possible mechanisms.

### Appendix A. Experiment instructions

After the subjects completed an experiment consent form and filled out demographic information, they were prompted with the following incentive information regarding payments for correct answers.

You will now be awarded a bonus to answer seven questions. The amount of bonus money you will earn in this HIT depends on your performance. To start, you have been awarded \$1.00. However, for **each** correct answer you will **gain** an additional \$0.25.

As an attention check, if you get one question correct, how much will you be awarded? (If you answer incorrectly the survey will end.)

- \$0.50
- \$1.25
- \$5.00

*Figure A.1. Incentive prompt for Gain treatment.*

You will now be awarded a bonus to answer seven questions. The amount of bonus money you will earn in this HIT depends on your performance. To start, you have been awarded \$2.75. However, for **each** incorrect answer you will **lose** \$0.25.

As an attention check, if you get **one** question incorrect, how much will you be awarded? (If you answer incorrectly the survey will end.)

- \$1.00
- \$2.50
- \$8.00

*Figure A.2. Incentive prompt for Loss treatment.*

You will now be awarded a bonus to answer seven questions. The amount of bonus money you will earn in this HIT does not depend on your performance. You will be awarded \$1.00 for completing the questions.

As an attention check, how much will you be awarded for completing the questions? (If you answer incorrectly the survey will end.)

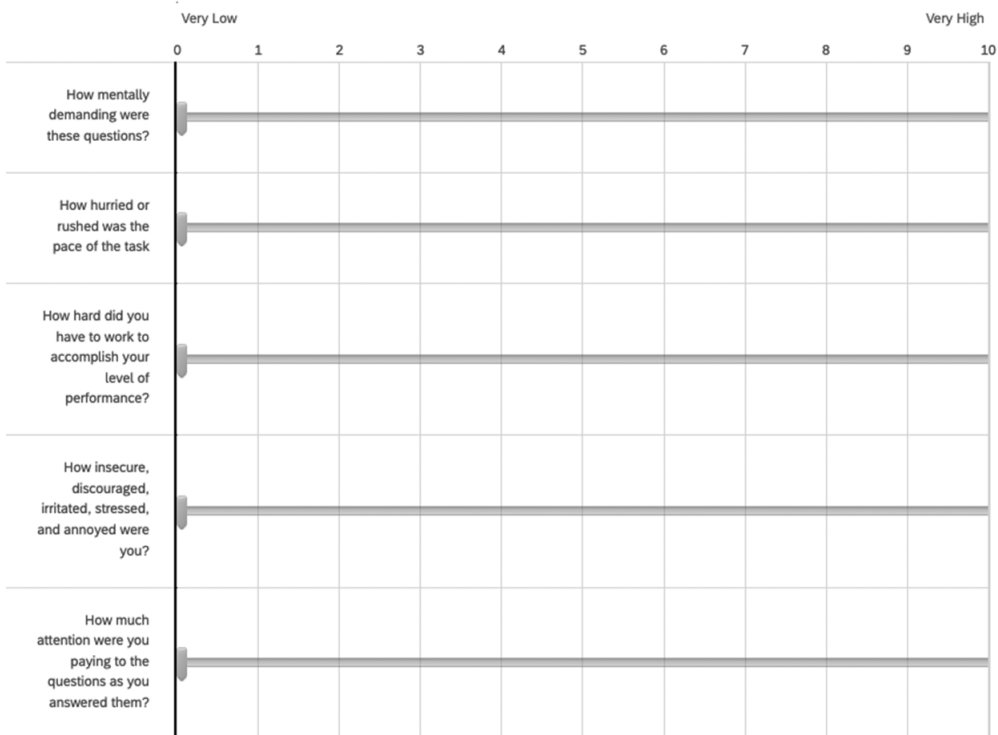
- \$0.50
- \$1.00
- \$2.00

**Figure A.3.** Incentive prompt for No Reward treatment.

After subjects received on of the prompts above, they were sequentially presented with the CRT questions in Table B.1 and the order of questions was the same for each subject.

After completing the CRT questions, the subjects were prompted with the following NASA TLX questions:

Considering the seven questions you just answered, please answer the following on a scale from 1 (very low) to 10 (very high):



**Figure A.4.** NASA TLX questions

**Appendix B. Cognitive reflection questions**

*Table B.1. CRT questions.*

| CRT questions  | Modified CRT questions  |
|--|---|
| <p>1 A bat and a ball cost \$1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost?<br/>__ cents</p> <p>CA: 5 cents, IA: 10 cents</p>  | <p>A rod and a reel cost \$3.50 in total. The rod costs three dollars more than the reel. How much does the reel cost?<br/>__ dollars</p> <p>CA: 0.25 dollars, IA: 0.5 dollars</p>  |
| <p>2 If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?<br/>__ minutes</p> <p>CA: 5 minutes, IA: 100 minutes</p>   | <p>If it takes 7 squirrels 7 minutes to collect 7 nuts, how long would it take 50 squirrels to collect 50 nuts?<br/>__ minutes</p> <p>CA: 7 minutes, IA: 50 minutes</p>   |
| <p>3 In a lake, there is a patch of lily pads. Everyday, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?<br/>__ days</p> <p>CA: 47 days, IA: 24 days</p>  | <p>There is a virus in the city you live. Every day, the number of people infected doubles in size. If it takes 64 days for the whole city to be infected how long would it take for half the city to be infected?<br/>__ days</p> <p>CA: 63 days, IA: 32 days</p>  |
| <p>4 If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together?<br/>__ days</p> <p>CA: 4 days, IA: 9 days</p>  | <p>If Jose can drink one bottle of Coke in 4 hours, and Beth can drink one bottle of Coke in 12 hours, how long would it take them to drink one bottle of Coke together?<br/>__ hours</p> <p>CA: 3 hours, IA: 8 hours</p>   |
| <p>5 Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?<br/>__ students</p> <p>CA: 29 students, IA: 30 students</p>   | <p>Michael finished both the 10th highest and the 10th lowest in the tournament. How many people are in the tournament?<br/>__ people</p> <p>CA: 19 people, IA: 20 people</p>   |
| <p>6 A man buys a pig for \$60, sells it for \$70, buys it back for \$80, and sells it finally for \$90. How much has he made?<br/>__ dollars</p> <p>CA: \$20, IA: \$10</p>  | <p>A woman buys a goat for \$30, sells it for \$40, buys it back for \$50, and sells it finally for \$60. How much has she made?<br/>__ dollars</p> <p>CA: \$20, IA: \$10</p>   |
| <p>7 Simon decided to invest \$8,000 in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Simon has:</p> <p>a. broken even in the stock market, b. is ahead of where he began, c. has lost money</p> <p>CA: c, IA: b</p> | <p>Deborah decided to invest \$5,000 in Bitcoin one day early in 2016. Five months after she invested, on August 27th, the value of her Bitcoin investment was down 50%. After this, from August 27th to November 27th, the Bitcoin she had purchased went up 75%. At this point, compared to her starting position, Deborah has:</p> <p>a. broken even in her Bitcoin investment, b. is ahead of where he began, c. has lost money</p> <p>CA: c, IA: b</p> |

Note: CA and IA denote “Correct Answer” and “Intuitive Answer,” respectively.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/jdm.2024.19>.

**Acknowledgments.** We thank seminar participants at the 2021 Liberal Arts Experimental Conference, the 2022 ESA meetings, and the 2022 New England Experimental Conference. Our study was pre-registered at the AEA RCT registry (ID:AEARCTR-0007839).

## References

- Alós-Ferrer, C. (2018). A dual-process diffusion model. *Journal of Behavioral Decision Making*, *31*, 203–218.
- Alós-Ferrer, C., & Garagnani, M. (2023). Part-time Bayesians: Incentives and behavioral heterogeneity in belief updating. *Management Science*, *69*, 5523–5542.
- Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*, 569.
- Amiez, C., Joseph, J.-P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *European Journal of Neuroscience*, *21*, 3447–3452.
- Arsalidou, M., Pascual-Leone, J., Johnson, J., Morris, D., & Taylor, M. J. (2013). A balancing act of the brain: Activations and deactivations driven by cognitive load. *Brain and Behavior*, *3*, 273–285.
- Bago, B. & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109.
- Bourgeois-Gironde, S. & Van Der Henst, J.-B. (2009). How to open the door to system 2: Debiasing the bat-and-ball problem. In S. Watanabe, A. P. Blaisdell, L. Huber, & A. Young (Eds.), *Rational animals, irrational humans*. Tokyo: Keio University, chap. 14.
- Brañas-Garza, P., Kujal, P., & Lenkei, B. (2019). Cognitive reflection test: Whom, how, when. *Journal of Behavioral and Experimental Economics*, *82*, 101455.
- Bronchetti, E. T., Kessler, J. B., Magenheimer, E. B., Taubinsky, D., & Zwick, E. (2023). Is Attention Produced Optimally? Theory and Evidence from Experiments With Bandwidth Enhancements. *Econometrica*, *92*, 669–707.
- Camerer, C. F. & Malmendier, U. (2007). Behavioral economics of organizations. In *Yrjö Jahnsson foundation 50th anniversary conference*, Jun, 2004, Espoo, Finland: This book is based on the aforementioned conference, Princeton University Press.
- Chater, N. (2018). Is the type 1/type 2 distinction important for behavioral policy? *Trends in Cognitive Sciences*, *22*, 369–371.
- Chen, X., Voets, S., Jenkinson, N., & Galea, J. M. (2020). Dopamine-dependent loss aversion during effort-based decision-making. *Journal of Neuroscience*, *40*, 661–670.
- Clay, S. N., Clithero, J. A., Harris, A. M., & Reed, C. L. (2017). Loss aversion reflects information accumulation, not bias: A drift-diffusion model study. *Frontiers in Psychology*, *8*, 1708.
- Cowell, P. E., Sluming, V. A., Wilkinson, I. D., Cezayirli, E., Romanowski, C. A., Webb, J. A., Keller, S. S., Mayes, A., & Roberts, N. (2007). Effects of sex and age on regional prefrontal brain volume in two human cohorts. *European Journal of Neuroscience*, *25*, 307–318.
- De Neys, W. (2023). Advancing theorizing about fast-and-slow thinking. *Behavioral and Brain Sciences*, *46*, e111.
- Deck, C., S. Jahedi, & R. Sheremeta (2021). On the consistency of cognitive load. *European Economic Review*, *134*, 103695.
- Dong, R., Fisman, R., Wang, Y., & Xu, N. (2021). Air pollution, affect, and forecasting bias: Evidence from Chinese financial analysts. *Journal of Financial Economics*, *139*, 971–984.
- Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.
- Evans, J. S. B. & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, *11*, 382–389.
- Evans, J. S. B. & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*, 223–241.
- Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*, 1–17.
- Foti, D., Weinberg, A., Bernat, E. M., & Proudfit, G. H. (2015). Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clinical Neurophysiology*, *126*, 1338–1347.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42.
- Frydman, C. & Camerer, C. F. (2016). The psychology and neuroscience of financial decision making. *Trends in Cognitive Sciences*, *20*, 661–675.
- Fujiwara, J., Tobler, P. N., Taira, M., Iijima, T., & Tsutsui, K.-I. (2009). Segregated and integrated coding of reward and punishment in the cingulate cortex. *Journal of Neurophysiology*, *101*, 3284–3293.
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279–2282.
- Gillard, E., Van Dooren, W., Schaeken, W., & Verschaffel, L. (2009). Proportional reasoning as a heuristic-based process: Time constraint and dual task considerations. *Experimental Psychology*, *56*, 92–99.
- Goel, V., Buchel, C., Frith, C., & Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage*, *12*, 504–514.
- Hochman, G., Glöckner, A., & Yechiam, E. (2009). Physiological measures in identifying decision strategies. *Foundations for tracing intuition*, 147–167. East Sussex: Psychology Press.



- Hochman, G. & Yechiam, E. (2011). Loss aversion in the eye and in the heart: The autonomic nervous system's responses to losses. *Journal of Behavioral Decision Making*, 24, 140–156.
- Hossain, T. & List, J. A. (2012). The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science*, 58, 2151–2167.
- Huang, N., Burtch, G., Gu, B., Hong, Y., Liang, C., Wang, K., Fu, D., & Yang, B. (2019). Motivating user-generated content with performance feedback: Evidence from randomized field experiments. *Management Science*, 65, 327–345.
- Isler, O. & Yilmaz, O. (2023). How to activate intuitive and reflective thinking in behavior research? A comprehensive examination of experimental techniques. *Behavior Research Methods*, 55, 3679–3698.
- James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt and Company.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Macmillan.
- Kaller, C. P., Heinze, K., Mader, I., Unterrainer, J. M., Rahm, B., Weiller, C., & Köstering, L. (2012). Linking planning performance and gray matter density in mid-dorsolateral prefrontal cortex: Moderating effects of age and sex. *Neuroimage*, 63, 1454–1463.
- Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review*, 23, 1–19.
- Kounios, J. & M. Beeman (2009). The Aha! moment: The cognitive neuroscience of insight. *Current Directions in Psychological Science*, 18, 210–216.
- Lejarraga, T., Schulte-Mecklenbeck, M., Pachur, T., & Hertwig, R. (2019). The attention–aversion gap: How allocation of attention relates to loss aversion. *Evolution and Human Behavior*, 40, 457–469.
- Malmendier, U. (2018). Behavioral corporate finance. In *Handbook of behavioral economics: Applications and foundations 1*. (Vol. 1, pp. 277–379). Amsterdam: North Holland Publishing Company.
- Melnikoff, D. E. & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences*, 22, 280–293.
- Morewedge, C. K. and Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, 14, 435–440.
- Oprea, R. (2020). What makes a rule complex? *American Economic Review*, 110, 3913–3951.
- Oprea, R. (2022). Simplicity equivalents. Working paper.
- Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences*, 9, 60–68.
- Sallet, J., Quilodran, R., Rothé, M., Vezoli, J., Joseph, J.-P., & Procyk, E. (2007). Expectations, gains, and losses in the anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience*, 7, 327–336.
- Shepherd, D. A., Williams, T. A., & Patzelt, H. (2015). Thinking about entrepreneurial decision making: Review and research agenda. *Journal of Management*, 41, 11–46.
- Siegel, G. & Ramanauskas-Marconi, H. (1989). *Behavioral accounting*. Mason, OH: Thomson South-Western.
- Smith, V. L. & Walker, J. M. (1993). Monetary rewards and decision cost in experimental economics. *Economic Inquiry*, 31, 245–261.
- Stango, V. & J. Zinman (2023). We are all behavioural, more, or less: A taxonomy of consumer decision-making. *The Review of Economic Studies*, 90, 1470–1498.
- Stanovich, K. E. & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127, 161.
- Stanovich, K. E. & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645–665.
- Thaler, R. H. (2005). *Advances in behavioral finance*, Vol. 2. Russell Sage Foundation.
- Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, 315, 515–518.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning*, 20, 147–168.
- Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the cognitive reflection test. *Cognition*, 150, 109–118.
- Van Hoof, J., Verschaffel, L., De Neys, W., & Van Dooren, W. (2020). Intuitive errors in learners' fraction understanding: A dual-process perspective on the natural number bias. *Memory & Cognition*, 48, 1171–1180.
- Wallis, J. D. & Kennerley, S. W. (2011) Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex. *Annals of the New York Academy of Sciences*, 1239, 33–42.
- Westbrook, A. & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective, & Behavioral Neuroscience*, 15, 395–415.
- Yechiam, E. & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin*, 139, 497.
- Yechiam, E. & Zeif, D. (2023). Revisiting the effect of incentivization on cognitive reflection: A meta-analysis. *Journal of Behavioral Decision Making*, 36, e2286.

**Cite this article:** Carpenter, J. and Munro, D. (2025). Do losses trigger deliberative reasoning? *Judgment and Decision Making*, e11. <https://doi.org/10.1017/jdm.2024.19>