2016

# The Effects of Experience, Choice Architecture, and Cognitive Reflection in Individual and Strategic Decisions

Mark Schneider
*Chapman University*

David Porter
*Chapman University*, dporter@chapman.edu



## Recommended Citation

Schneider, M., & Porter, D. (2016). The effects of experience, choice architecture, and cognitive reflection in individual and strategic decisions. *ESI Working Paper 16-24*. Retrieved from http://digitalcommons.chapman.edu/esi_working_papers/202

# The Effects of Experience, Choice Architecture, and Cognitive Reflection in Individual and Strategic Decisions

## Comments

Working Paper 16-24

This paper was later published as:

Schneider, M., & Porter, D. (2020). Effects of experience, choice architecture, and cognitive reflection in strategyproof mechanisms. *Journal of Economic Behavior & Organization*, *171*, 361–377. https://doi.org/10.1016/j.jebo.2020.01.012

**The Effects of Experience, Choice Architecture, and Cognitive Reflection in Individual and Strategic Decisions**

Mark Schneider* David Porter*

**This Version: January 18, 2018**

**Abstract**

We study how performance in strategyproof mechanisms and individual lottery choices is affected by experience with the decision environment, choice architecture (selection among strategically equivalent mechanisms), and cognitive reflection. In both individual and strategic decisions, we observe substantial gaps in performance between high reflective and low reflective participants. We also find that choice architecture and experience narrow these gaps in performance. Our primary finding is that experience serves as a substitute for cognitive reflection: Across a series of experiments employing multiple rounds of a lottery task, a second price sealed bid auction, an English clock auction, and a random serial dictatorship allocation mechanism, we consistently find that the performance of low reflection participants with experience is similar to that of high reflection participants without experience. We also find across all tasks that switching from a strategyproof to an 'obviously strategyproof' mechanism has a larger effect on performance than the difference of having a low level versus a high level of cognitive reflection, providing evidence that choice architectures can systematically induce or reduce the prevalence of rational behavior. A policy implication emerging from our results is that transparent mechanisms and familiar mechanisms (those with which participants have experience) can serve to increase the frequency of optimal decisions and efficient allocations in society.

Keywords: Cognitive Reflection; Stochastic Dominance; Second Price Auction; English Clock Auction

**JEL Codes: C91, C92, D01**

**Acknowledgements:** The authors thank Jeffrey Kirchner for his help with the design of the auction and lottery software, Shengwu Li for sharing his software for the random serial dictatorship mechanisms, and the Economic Science Institute for financial support.

*Economic Science Institute, Chapman University. One University Drive, Orange, CA 92866

## 1. Introduction

In recent years interest has grown in allocation procedures or ‘mechanisms’ that are dominant strategy incentive compatible. In such mechanisms, it is optimal, regardless of one’s beliefs about other agents’ actions, to report one’s true preferences to the mechanism. Mechanisms with this property are ‘strategyproof’: agents cannot improve their payoffs by strategically misreporting their preferences. In many cases, there are multiple ways to implement a given dominant strategy equilibrium. For instance, implementing a second price sealed bid auction or an English clock auction should, in theory, implement the same equilibrium outcome (Vickrey, 1961). However, in order for players to play a dominant strategy, they must first know that they have a dominant strategy and they must know which strategy is dominant. Recognizing this, Li (2017) formalized a refinement of dominant strategy incentive compatible mechanisms called ‘obviously strategyproof’ (OSP) mechanisms in which it is easier to identify the dominant strategy than in mechanisms which are merely strategyproof but not OSP.

Decisions with dominant strategies have not traditionally been considered as a source of heterogeneity in behavior. Even those who doubt that people routinely conform to the independence axiom of expected utility theory might still expect people to select a dominant strategy when it is available. This distinction is naturally captured by a distinction between two forms of rational behavior proposed by Gilboa et al. (2010) which they refer to as ‘objective rationality’ and ‘subjective rationality’. They argue that a choice is ‘objectively rational’ if the decision maker can convince others she made the right choice. In contrast, they refer to a choice as 'subjectively rational' if the decision maker cannot be convinced she made the wrong choice. Under these qualitative definitions, subjective rationality pertains to matters of preference and belief (since there is ‘no disputing tastes’). A decision maker whose decisions are consistent with her own preferences and who best-responds to her beliefs about the actions of others conforms to subjective rationality. Objective rationality naturally pertains to decisions and games with dominant strategies.

Much research has focused on measuring the parameters of subjective rationality such as risk preferences, time preferences, and subjective probabilities. Considerably less attention has been devoted to identifying differences in objective rationality, partially because all agents are traditionally assumed to play dominant strategies when available, and thereby all agents have the same degree of objective rationality. Moreover, while there are standard methods for eliciting risk and time preferences, it is not clear what would be a reliable method for identifying heterogeneity in objective rationality in the narrowly defined sense of choosing dominant strategies when such strategies are available.

As a simple a priori measure of objective rationality for both individual and strategic settings, we employ a standard test for identifying a person’s natural tendency to reflect on his or her thought processes. Frederick (2005) introduced a three-question ‘cognitive reflection’ test (CRT) where the questions have intuitive but wrong answers and correct answers which require some reflection and found that it correlates with risk and time preferences. Subsequent research has employed the CRT and related measures of cognitive skills to identify a propensity to engage in backward induction (e.g., Burnham et al., 2009; Levitt et al., 2011; Branas-Garza et al., 2012), and the ability of a market of traders who have all high levels or have all low levels of cognitive reflection to aggregate information (Corgnet et al., 2015).

Using objective rationality in the narrowly defined sense of conforming to dominant strategies, we investigate whether the CRT can sort out differences in objective rationality across six mechanisms: two auctions (a second price sealed bid auction and an English clock auction), two random priority allocation mechanisms, and two choice architectures for making individual choices between lotteries. Using the same parameter values within each pair of mechanisms, each pair should, in principle, implement the same dominant strategy equilibrium. However, one mechanism in each pair is 'obviously strategyproof' (OSP) as defined by Li (2017) while the other is merely strategyproof[1] (SP). A boundedly rational agent may be more likely to recognize the dominant strategy in an OSP mechanism, than in an SP mechanism. The tasks in the experiment are summarized in Table 1.

**Table 1. Overview of Tasks in the Experiment**

| | **Games of Incomplete Information** | |
|---|---|---|
| | Non-Transparent Mechanism | Transparent Mechanism |
| Population | Second Price Sealed Bid Auction | English Clock Auction |
| Low CRT | 3 sessions / 2 iterations | 3 session / 2 iterations |
| High CRT | 3 sessions / 2 iterations | 3 session / 2 iterations |
| | **Games of Complete Information** | |
| | Non-Transparent Mechanism | Transparent Mechanism |
| Population | Static Serial Dictatorship | Dynamic Serial Dictatorship |
| Low CRT | 4 sessions / 2 iterations | 2 sessions / 2 iterations |
| High CRT | 4 sessions / 2 iterations | 2 sessions / 2 iterations |
| | **Individual Choices under Risk** | |
| Population | Non-Transparent Frame | Transparent Frame |
| Low CRT | 4 sessions / 2 iterations | 2 sessions / 2 iterations |
| High CRT | 4 sessions / 2 iterations | 2 sessions / 2 iterations |

In each session, we implemented a mechanism for two iterations to observe the effects of experience. The serial dictatorship mechanism and lottery choice tasks were conducted within the same experimental session in either a 'transparent session' (employing the dynamic serial dictatorship and the

[1] We use SP to denote the class of mechanisms that are strategyproof but not obviously strategyproof.

transparent lottery frame) or in a 'non-transparent session' (employing the static serial dictatorship and the non-transparent lottery frame). The order of tasks was counter-balanced such that the non-transparent lottery task was conducted first in two of the four non-transparent sessions and it was conducted second in the remaining two non-transparent sessions. The order of tasks in transparent sessions was likewise counterbalanced.

In our experiment, subjects make repeated decisions in each mechanism, enabling us to observe the effects of experience. By comparing the performance of the mechanisms within each pair, we can examine a role for choice architecture (which in our context involves selecting among strategically equivalent strategyproof mechanisms) to improve welfare. Our focus is on how the effects of cognitive reflection are moderated by experience and choice architecture. Our main findings are:

(i) Cognitive reflection is a reliable measure of objective rationality in novel SP decision environments.

(ii) Experience and choice architecture (using OSP instead of SP mechanisms) *each reduce* the gap in objective rationality between high and low reflection subjects and *jointly eliminate* the gap.

(iii) For all six mechanisms, experience serves as a substitute for cognitive reflection: Low reflection subjects with experience perform similarly to high reflection subjects without experience.

(iv) For all iterations in all treatments, the effect of transparency is greater than the effect of cognitive reflection.

(v) For the individual choices between lotteries, we find that surprisingly little feedback is needed to substantially increase the percentage of optimal decisions, even when the expected gains from choosing optimally are small.

To test for violations of objective rationality, the mechanisms used in our experiment were selected to each have a dominant strategy. In particular, we conducted experimental second price sealed bid auctions based on the design of Kagel and Levin (1993), separately for groups of high CRT subjects and low CRT subjects. In such auctions, there is a dominant strategy for how a person should bid, although previous experiments have found that many subjects do not 'discover' this strategy (Kagel and Levin, 1993; Kagel et al., 1987; Kagel, 1995). The English clock auction we employed was designed analogously to the second price design but where bidders can choose to exit the auction by clicking a button rather than clicking a button to submit their sealed bids. The two random priority allocation mechanisms (a static and a dynamic random serial dictatorship) were the same as in Li (2017). The two choice architectures for deciding between pairs of lotteries were similar to an example from Tversky and Kahneman (1986). However, none of these mechanisms have been studied in conjunction with cognitive reflection.

The remainder of this paper is organized as follows: We provide further background and motivation in Section 2. We then describe the experimental design and results for the auctions (Section 3), the design and results for the random priority mechanisms (Section 4), and the design and results for the lottery choices (Section 5). Section 6 concludes. The instructions for all mechanisms are in the supplementary material.

## 2. Background and Motivation

Several papers have considered the role of cognitive reflection in strategic settings that have dominated strategies with a focus on guessing games (also referred to as ‘p’-beauty contests) in particular. For instance, Burnham et al. (2009) and Branas-Garza et al. (2012) observe that students with lower CRT scores are more likely to play dominated strategies. Schnusenberg and Gallo (2011) replicate this finding but also observe that low CRT subjects make lower guesses in later rounds when the game is repeated.

Our focus is on games and decisions where there is a dominant strategy. Li (2017) experimentally studied a second price auction, an English auction, and two random priority allocation mechanisms in order to compare OSP and SP mechanisms. Zhang and Levin (2017) conducted a related experiment with a random serial dictatorship and individual choice task. However, Li (2017) and Zhang and Levin (2017) did not test for the role of cognitive reflection in explaining heterogeneity in the behavior of their subjects. One aspect of our study is then to investigate whether differences in cognitive reflection can sort out heterogeneity in bidding behavior in second price sealed bid auctions and in English clock auctions, and how cognitive reflection might affect revenue and efficiency in these auctions. A second aspect of our study is to consider the six mechanisms collectively and ask broader questions of whether cognitive reflection sorts out heterogeneity in objectively rational choices and how such a relationship might be moderated by experience with a mechanism and by choice architecture (e.g., redesigning a SP mechanism into an OSP mechanism).

## 3. Games with Incomplete Information

Our main experiment tests the effects of cognitive reflection, experience, and choice architecture across two classical mechanisms – a second price sealed bid auction and an English clock auction.

### 3.1 Experimental Design for the Second Price Auction

Eighty one[2] undergraduate students at a private California university participated in the second price auction experiment. Six sessions were conducted – three “High CRT” sessions and three “Low CRT” sessions. Subjects in each of these sessions had previously taken the seven-question Cognitive Reflection Test (Toplak et al., 2014) which is an extension of the original three-question cognitive reflection test (CRT) due to Frederick (2005), at an earlier date when they signed up to participate in economic experiments. Knowing the subjects’ CRT scores before they come to participate in an experiment makes it possible to

[2]Eighty-four subjects were recruited for the experiment (six experimental sessions with the lab’s capacity of fourteen subjects per session). The high CRT population is smaller than the low CRT population. While the first two high CRT second price sessions each had 14 subjects, only eleven high CRT subjects attended our last high CRT session for the second price auction.

recruit subjects who only obtained particular scores (e.g., high scores or low scores). We hypothesized that sampling from the tails of the distribution would reveal the starkest difference in performance based on CRT scores and reduce the noise in the measurement of a subject's tendency to cognitively reflect. We thus recruited "Low CRT" auction sessions in which all subjects had previously scored 0 or 1 on the CRT, as well as "High CRT" auction sessions in which all subjects had previously scored in the top 20% of the distribution of CRT scores (subjects who scored a 5, 6, or 7 on the CRT).

The second price auction experiments were based on the design of Kagel and Levin (1993). In each auction period, subjects participated in both a large market (where subjects competed in a group of 10 bidders) and a small market (where subjects competed in a group of 5 bidders)[3], by submitting a bid in each market via their bidding dashboard. Each subject received the same private valuation in the large market and the small market, but private valuations differed across subjects and across auction periods. For each auction period, private valuations were randomly drawn from a discrete uniform distribution with step size of $0.01, over the interval [$0.00, $28.30] which was the same distribution employed in Kagel and Levin (1993). Subjects knew their private valuation, the distribution from which all values were drawn, and the total number of bidders in each market. Subjects did not need to recall this information as it was always displayed to them on their market dashboard, as shown in Figure 1. After each period, the dashboard also displayed the winning bid, the profit made by the winning bidder and the subject's own bid. In each period, either the large market or the small market was randomly selected for payment. As in Kagel and Levin (1993), subjects were each given a starting cash balance of $10 to cover the possibility of losses.

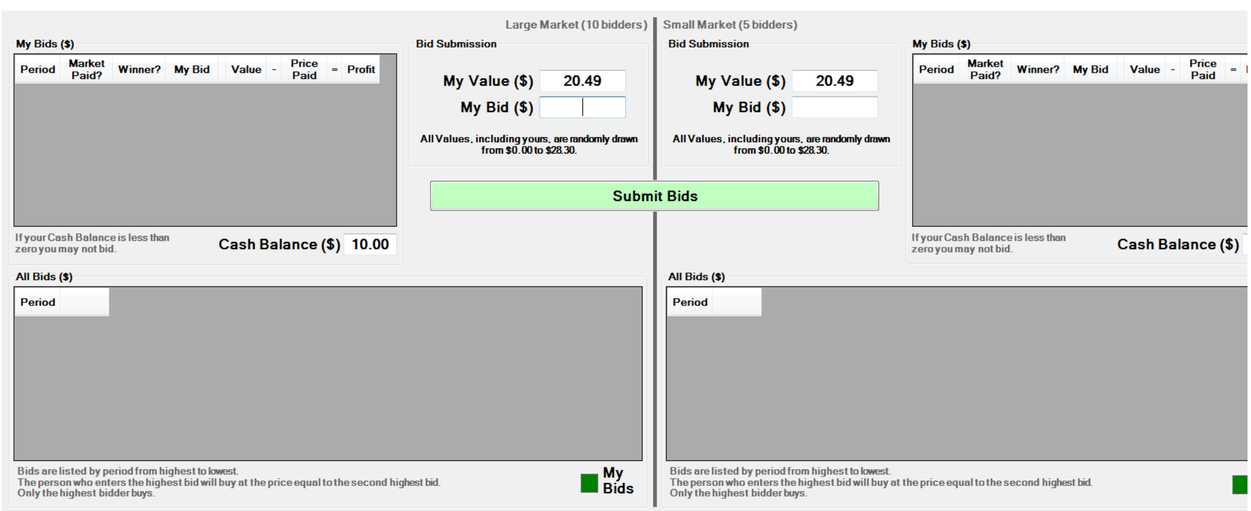



**Figure 1. Market Dashboard from Second Price Auction.** Each subject submits a bid in a large market (10 bidders) and a small market (5 bidders) in each period, with the same valuation in both markets.

[3]In auction periods where more than the number of 'reserve' bidders had gone bankrupt (i.e., their cash balance had gone negative), the large (small) market contained less than 10 (5) bidders.

We conducted six experimental second price auction sessions, three each for high CRT subjects and low CRT subjects. A total of 81 subjects (39 high CRT subjects and 42 low CRT subjects) participated in one of the auction sessions. In each session, each subject was seated at a separate computer terminal in a cubicle such that no subject could observe the actions or computers of other subjects. Each auction session (three each for high CRT and low CRT subjects) involved two iterations of 20 rounds each. That is, subjects participated in 20 rounds of the second price auction and their earnings were calculated. Gains or losses in each period were added to each subject's balance. If a subjects' balance went negative, they were no longer permitted to bid in that auction iteration. At the beginning of the experiment, subjects were informed that they would participate in two iterations of the experiment, that they would be paid the sum of their earnings across both iterations, and that their balance would be reset to $10 before the second iteration (so any losses did not carry over). Conducting two iterations in a session enabled us to investigate potential experience effects. Since many low CRT subjects went bankrupt in the first iteration (i.e., their cash balance went negative), conducting a second iteration also enabled us to observe the outcome of a full session of active bidders, as very few low CRT subjects went bankrupt twice.

At the start of each auction session, the large market contained ten bidders and the small market contained five bidders. It was intended for the large market and the small market to retain their respective sizes across all auction periods but this was not always possible in later auction periods due to bankruptcies. To anticipate this possibility, following Kagel and Levin, we recruited more than ten subjects and each subject was randomly assigned to 'play' or 'observe' in each period. Doing so allows for 'reserve bidders' to maintain the size of the large and small markets in case of bankruptcies. There were typically four extra bidders in each auction session, but even this was not always sufficient to keep the number of bidders constant in the large and small markets. The software was programmed with a schedule of how to adjust the market sizes in the case of bankruptcies. The large market contained all remaining bidders when the total number of bidders dropped below ten. The small markets were balanced to be as close in size as possible.

Subjects were given detailed instructions, which are provided in the electronic supplementary material. They were informed of the second price rule for selecting the winning bidder and how payments were determined. Subjects were also informed that they could not bid more than $50 for the item being auctioned. No statements were included which could be seen as censoring the bidding process or nudging bidders in a certain direction such as "It is possible to lose money if you bid above your value, but not if you bid below your value." Rather, after explaining the rules, we wanted to provide as little nudging as possible to give bidders the opportunity to discover the dominant bidding strategy without any 'hints'.

Viewing interactive learning to also be effective in helping participants understand the rules of the auction, each participant saw three interactive examples, one each in which they were assigned a low value, an intermediate value, and a high value. In each example participants submitted bids in their bidding dashboard and computerized agents were programmed with a fixed set of bids to complete the auction. From this part of the instructions, subjects could experience the bidding process and observe their profits or losses at no cost to themselves. Subjects were also quizzed by the software on the auction instructions and were paid $0.50 for each correct answer they provided to the five-question quiz. After all subjects completed the instructions, the experiment began. After all auction periods had ended, subjects were paid their earnings from the quiz and from the auction periods in cash in addition to a $7 participation payment.

### 3.2 Experimental Design for the English Clock Auction

To study the relationship between cognitive reflection, experience, and mechanism transparency, we also conducted six ‘English Clock Auction’ sessions with a different sample of undergraduate students from the same California university as in the second price experiment. Eighty three[4] undergraduate students (41 who scored either a 0 or 1 on the seven-question CRT and 42 who scored either 5, 6, or 7) participated in the second price auction experiment. Six sessions were conducted – three “High CRT” sessions consisting only of participants with a score of 5, 6, or 7 on the CRT, and three “Low CRT” sessions, consisting only of participants with a 0 or 1 on the CRT. In each session, subjects participated in two iterations of the English clock auction. The first iteration of the clock auction consisted of 20 periods as in the second price sessions. Since the clock auctions take much longer than the second price sealed bid auctions (fixing the clock at a constant tick rate), and since subjects participated in both a large and small market clock auction to parallel the second price sessions, there was not sufficient time to also conduct 20 periods for the second iteration of the clock auction in a two-hour experiment. Instead, the second iteration of the clock auction contained 10 periods, which we believed would be sufficient to study how behavior differs with experience, with CRT, and between the second price and clock auctions. Subjects thus participated in sixty clock auctions – one large clock auction (10 bidders) and one small clock auction (5 bidders) in each of 30 periods.

For all periods of the clock auction, the value distribution and the particular set of values in each auction period as well as the sequence of periods in which these values occurred was exactly the same as in the second price sessions so that each period in the English clock auction had the same dominant strategy equilibrium as the corresponding period in the second price auction.

In the clock auction periods, the large market always occurred first. This was done to be consistent with our focus on the large market in the second price session. The small market occurred immediately after bidding ended for the large market in each period. As in the second price sessions, subjects had the same value for the large and small markets. Whereas subjects entered a bid and clicked a “Submit Bid” button in

[4]Eighty-four subjects were recruited for the experiment (six experimental sessions with the lab’s capacity of fourteen subjects per session). One subject did not show.

the second price sealed bid auction, subjects in the clock auction clicked a “Do Not Buy” button when they no longer wanted to participate in that auction in each period. The last bidder to click the “Do Not Buy” button in an auction sets the price for that auction with the winning bidder being the subject who remains after all other bidders click “Do Not Buy.” An image of the ‘Large Market’ side of the Market Dashboard from the English Clock auction experiment is displayed in Figure 2.

In the English Clock auctions, subjects saw their value and the price of the ‘fictitious commodity’ being auctioned as it increased at increments of $0.25, starting at $0.00. Subjects were also shown their ‘potential profit’ (the difference between their value and the price) if the auction were to end at the current price. While such information is natural for an English Clock auction, it cannot be calculated for the second price sealed bid auction where there is no ‘current price’ prior to the close of the auction. As in the second price auction sessions, subjects were randomly paid for either the large market or the small market in each period. Subjects were paid their total earnings across all auction periods. This amount included a $10 starting cash balance from both the first and second iterations plus any gains or losses they incurred during the auction periods. Subjects could not lose more than their total earnings. In addition, subjects received $0.50 for each correct answer they provided to the quiz questions in the instructions. The instructions also included three interactive examples similar to those used in the second price instructions. All experimental instructions are contained in the supplementary material.

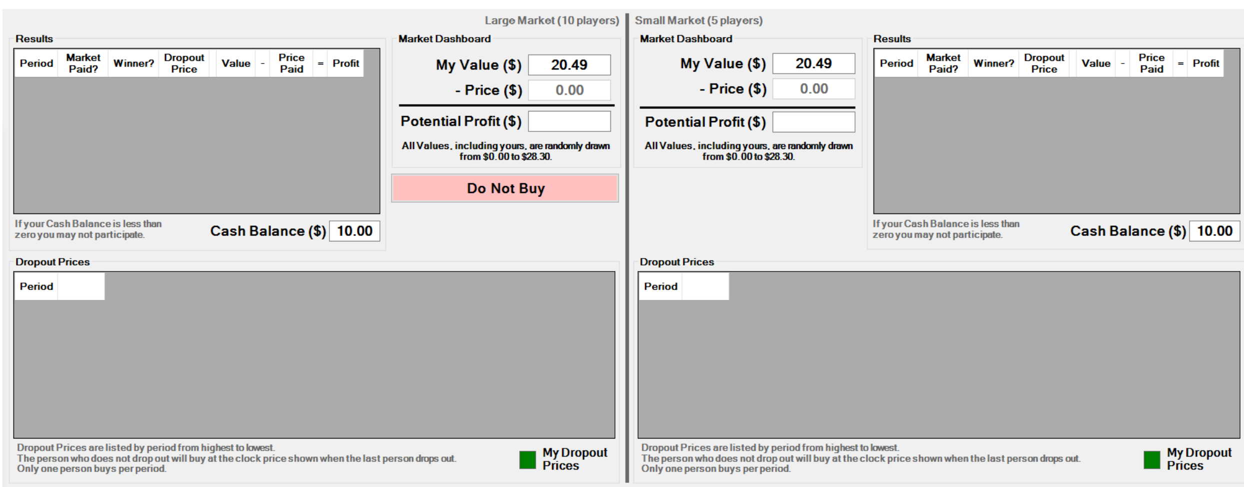



**Figure 2. Market Dashboard from English Clock Auction.** Each subject submits a bid in a large market (10 bidders) and a small market (5 bidders) in each period, with the same valuation in both markets.

### 3.3 Experimental Results

Since the classic work of Vickrey (1961), the second price auction (SPA) has attracted much attention in economics research due to its appealing properties. For instance, in such an auction with private valuations, it is a dominant strategy to bid exactly one’s valuation. That is, regardless of what other bidders do, it is optimal to bid your value. This dominant strategy equilibrium is a stronger property than a Nash equilibrium where one typically needs to invoke common knowledge assumptions about others' payoffs and their rationality and condition one's bidding strategy on how he expects others to bid.

To get a sense of our second price data without experience effects, we first compared the distribution of initial bids for high and low CRT subjects. In particular, we looked at the first bid made by each bidder across all experimental sessions, and computed (i) the average absolute difference (in dollars) between that bidder's bid and value, and (ii) the proportion of bidders who bid within $1 of their value on their initial bid. We performed these calculations for both the large and small markets. The average absolute deviation of bids from values for low CRT subjects in the large (small) market was $8.22 ($7.66). The average absolute deviation for high CRT subjects in the large (small) market was $2.01 ($1.94). The proportion of low CRT scorers who bid within $1 of their value on their first bid in the large (small) market was 0.214 (0.238). The proportion for high CRT scorers in the large (small) market was 0.590 (0.615). Thus, high CRT subjects were much more likely to bid within $1 of their value than low CRT subjects. The difference between the proportion of first bids within $1 of the value for high and low CRT subjects is significant for both the large and small markets (2-tailed Z difference in proportions test, $p = 0.00056$ for large market and $p = 0.00058$ for small market).

For an English auction, we cannot determine the bid of the winning bidder as that bidder remains active after all other bidders have dropped out. We can compute the average absolute deviation of bids from values for the initial bids of 'non-winning bidders' (those who did not win the first auction they participated in) as well as the proportion of these bidders who bid within $1 of their value on their first bid. The average absolute deviation of bids from values (for the initial bids of non-winning bidders) in the large market was $4.02 for low CRT subjects and $2.22 for high CRT subjects. The proportion of non-winning bidders who bid within $1 of their value on their initial bid in the larger market was 0.500 for low CRT bidders and was 0.684 for high CRT bidders. The results for the small market are similar. These statistics are not directly comparable to the second price auction data noted above since those results include the initial bids of all bidders (including those who won the auction with their initial bid). To more directly compare the initial bids for the English and second price auction experiments, we can look at the bidders in the second price auction who did not win an auction on their first bid. Doing so, we find that for the initial bids of non-winning bidders, the average absolute deviation of bids from values in the large market was $6.94 for low CRT subjects and $2.00 for high CRT subjects. In addition, the proportion of non-winning initial bids within $1 of the bidder's value is 0.237 for low CRT subjects and 0.556 for high CRT subjects. This difference is also significant (two-tailed Z difference in proportions test, $p < 0.01$). In contrast the difference between the corresponding proportions for the English auction (0.500 for low CRT subjects and 0.684 for high CRT subjects) is not significant.

Table 2 provides summary statistics for the second price and English clock sessions. The table displays (i) the number of subjects in each treatment, (ii) the proportion of subjects in these sessions who went bankrupt, (iii) the proportion of subjects who lost money relative to their $10 endowment, (iv) the average surplus (the average difference between the winning bidder's value and the second highest bid) per auction[5], (v) the proportion of efficient allocations across all auction periods for each iteration and (vi) the proportion of subjects whose average bias in a given iteration is within $1 of valuations in the large market.[6]

**Table 2. Data for Second Price Sealed Bid (SB) and English Clock (EC) Auctions in Large Market**

| Treatment | Iteration | N | Bankrupt | Lost Money | Average Surplus per auction ($) | Efficiency | Average Bias < $1 |
|---|---|---|---|---|---|---|---|
| SB Low CRT | First | 42 | 0.667 | 0.690 | -11.140 | 0.500 | 0.095 |
| SB Low CRT | Second | 42 | 0.119 | 0.238 | -1.540 | 0.600 | 0.381 |
| SB High CRT | First | 39 | 0.077 | 0.231 | -0.350 | 0.650 | 0.462 |
| SB High CRT | Second | 39 | 0.000 | 0.179 | 1.670 | 0.817 | 0.769 |
| EC Low CRT | First | 41 | 0.000 | 0.024 | 3.017 | 0.650 | 0.390 |
| EC Low CRT | Second | 41 | 0.000 | 0.000 | 3.703 | 0.800 | 0.610 |
| EC High CRT | First | 42 | 0.000 | 0.024 | 3.010 | 0.767 | 0.643 |
| EC High CRT | Second | 42 | 0.000 | 0.024 | 3.089 | 0.800 | 0.667 |

In Table 2, we observe large differences between low and high CRT subjects in the first iteration of the second price auction. For instance, roughly two-thirds of low CRT subjects went bankrupt indicating they would have been better off by not bidding at all (and walking away with their full $10 endowment), whereas less than 8% of high CRT subjects did so. Low CRT subjects averaged more than a $1 loss in both iterations of the second price auction, while high CRT subjects lost less than $1 on average in the first iteration and earned more than $1 in profit on average in the second iteration. In addition, less than 10% of low CRT subjects bid within $1 of their value on average in the first iteration of the second price auction, while 46% of high CRT subjects did so. The large number of bankruptcies in the second price auction resulted from persistent overbidding. Kagel and Levin (1993) also observed frequent overbidding in second price auctions. They note that earlier reports of convergence to the dominant strategy in second price auctions (Cox et al., 1982) employed procedures which prohibited bidding above valuations.

For the English Clock auction, the differences between low CRT and high CRT subjects is less stark. As already noted for the English auction, the proportion of low CRT subjects bidding within $1 of their value on their first bid did not differ significantly from high CRT subjects, while the difference was highly significant for the second price auction. From Table 2, we further see that the gap in average surplus per

[5]Whenever the difference between the winning bidder's value and the price (second highest bid) were greater than the winning bidder's cash balance, the winning bidder went bankrupt and could not bid in later periods.
[6]The average bias for a subject is calculated as the average absolute deviation of a subject's bid from that subject's value across all periods in which that subject was an active bidder. Statistics for the small markets are similar to those in Table 2.

auction between low and high CRT subjects is eliminated in the English auction. Still, there are some metrics on which high CRT subjects continue to perform better – particularly in the proportion of efficient allocations and in the proportion of subjects with an average bias of less than $1.

Our main finding from Table 2 is that experience serves as a substitute for cognitive reflection: low CRT subjects in their second iteration perform similarly to high CRT subjects in their first iteration. For instance, the proportion of subjects who went bankrupt or lost money, and the average earnings and efficient allocations are similar for 'experienced' low CRT subjects and 'inexperienced' high CRT subjects in both the second price and the English auction. This suggests that experience can narrow the gap in performance between high and low CRT subjects, thereby helping to compensate for differences in cognitive reflection. In the second price auction, for instance, it requires reflection *or experience* to recognize that bidding above your value can result in losses.

One other finding from Table 2 is that the transparency of a mechanism can narrow the gap between high and low CRT subjects. In particular, an English Clock auction is more transparent than a second price sealed bid auction since an English Clock auction makes the relationship between one's value and the price of the item easy to recognize: One need only compare the value and the price to decide when to drop out of the bidding. However, in a second price sealed bid auction, one has to select an amount to bid from a large message space, and it requires more contemplation to identify the dominant strategy.

Indeed, from Table 2, we see that low CRT subjects perform closer to high CRT subjects in the English Clock auction than in the second price sealed bid auction and both groups earn more money and generate a greater proportion of efficient allocations in the English auction.

To better visualize the difference between high CRT and low CRT subjects in the second price auction, Figure 3 displays the distribution of all bids in the large market in periods 1 through 10 of each session (a total of 100 bids for a session if there were no bankruptcies). Since many low CRT subjects went bankrupt in the first iteration before period 10, periods 11 through 20 would be biased, displaying the bids of only a few bidders who 'survived' the market.

From Figure 3 we can see that the high CRT subjects bid very close to their value, even in the first iteration, with relatively small deviations from truthful bidding. In contrast, low CRT subjects had a much wider and more volatile distribution of bids, deviating considerably from the dominant strategy equilibrium in which bids equal values. We can also see that low CRT subjects bid much closer to their values in the second iteration, relative to their first iteration.

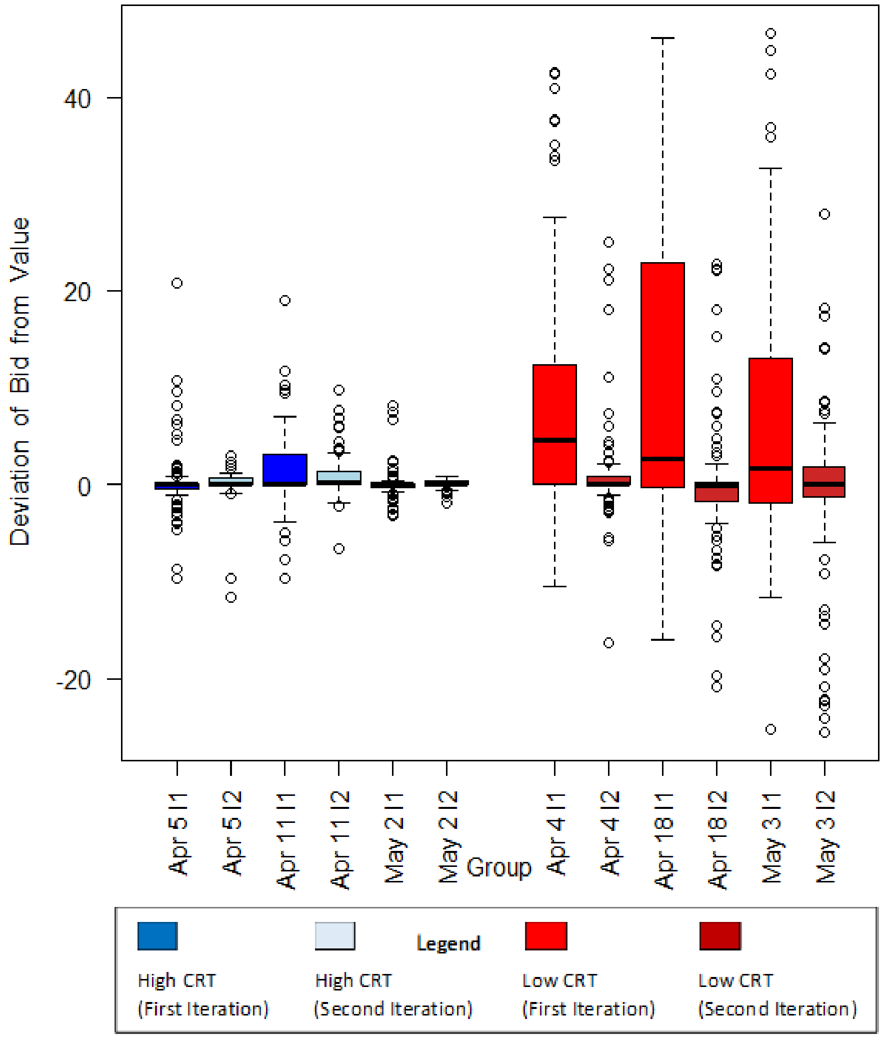
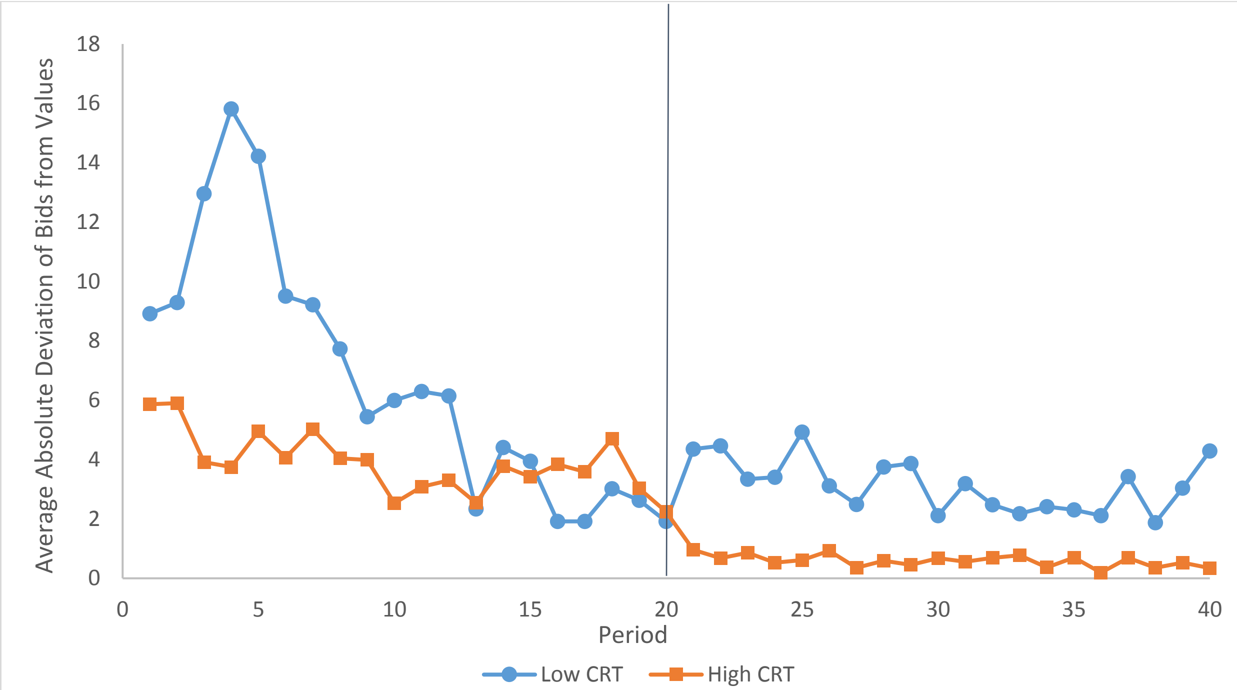



**Figure 3. Deviation of bids from values in the first and second iteration of the second price auction (large market, periods 1 through 10) for high CRT subjects (left) and low CRT subjects (right).** The line in the interior of each boxplot is the median deviation from bidding one's value in periods 1 through 10. The ends of each box display the first and third quartiles of the distribution. The ends of the whiskers extending from each box correspond to 1.5 times the interquartile range. Boxplots ending with "I1" ("I2") correspond to the first (second) iteration in a session. For instance, "Apr5I2" denotes the second iteration on April 5.

To observe the change in behavior across both iterations (40 periods), the average absolute deviation of bids from values among all active bidders in the large market is plotted in Figure 4 (with a vertical line dividing the data for Iterations 1 and 2) for both the low and high CRT groups. From the figure, we see that high CRT subjects converged closely to the dominant strategy equilibrium by the second iteration, while low CRT subjects approached but did not reach the equilibrium strategy by period 40. It is surprising that a simple seven-question test such as the CRT that bears no direct relationship to auctions or strategic behavior can accurately sort out how subjects are going to bid in second price sealed bid auctions. Finally, note that we again observe low CRT subjects with experience (see low CRT data points in Figure 4 for periods 21-40) perform similarly to high CRT subjects without experience (see high CRT data points in Figure 4 for periods 1 – 20).

**Figure 4. Average Absolute Deviation of Bids from Values among Active Bidders in Large Market**

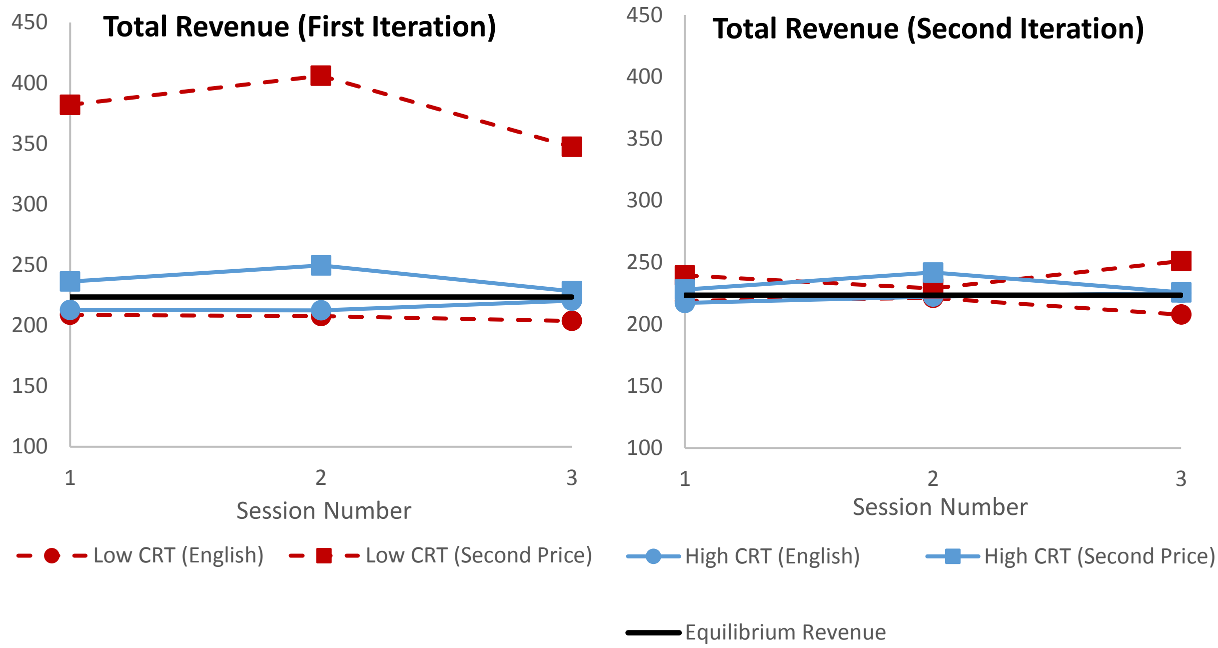



Although the high CRT auctions produced greater efficiency, the low CRT auctions produced greater revenue. Figure 5 displays the total revenue for periods 1 through 10 for low CRT second price and English auctions, high CRT second price and English auctions, and for the dominant strategy equilibrium. From Figure 5, we see that revenue from all second price auctions was higher than the equilibrium revenue, while revenues for the English auctions were slightly lower than the equilibrium revenue.

On average, the revenue from the first iteration of the second price auction for the low CRT sessions summing over periods 1 through 10 is $154.92 higher than the revenue if all players played their dominant strategies. In contrast, the revenue from the first iteration for the high CRT sessions summing over periods 1 through 10 is $14.49 higher than the equilibrium revenue. Low CRT subjects with experience again look similar to high CRT subjects without experience. In their second iteration, low CRT subjects pay $16.33 more than the equilibrium revenue summing over periods 1 through 10 (compared to $14.49 for high CRT subjects without experience). With experience, high CRT subjects pay $8.39 more than the equilibrium revenue. Experience thus considerably reduces deviations from equilibrium revenue for both groups.

For the English Clock auction, low CRT subjects with experience bid $7.52 less than the equilibrium revenue, whereas high CRT subjects without experience bid $8.27 less than the equilibrium revenue. High CRT subjects with experience bid within $2 of the equilibrium revenue.

**Figure 5. Low CRT, High CRT, and Equilibrium Revenue for First Iteration (left) and Second Iteration (right) from the Second Price and English Clock Auctions (Total Revenue across Periods 1 through 10).**




### 3.4 Regression Model

To further analyze the data, we implemented the following regression model with session fixed effects:

(1) $$ln\ (v_{ps}) = \alpha + \beta_H ln(\overline{v}_{pH}) + \beta_L ln(\overline{v}_{pL}) + \gamma x_{Cps} + \lambda x_{Tps} + \phi x_{CTps} + \textstyle\sum_s \kappa_s x_{ps} + \varepsilon_{ps}$$

where the dependent variable, $v_{ps}$ is the value of the winning bidder in period $p$ of session $s$. Log transformations were used so that the values can span the real line. The constant $\alpha$ is the mean effect size across all treatments and the other explanatory variables are parametrized as deviations from this mean.

Regression model (1) was implemented separately for the first iteration and the second iteration of the auction experiments. Only the first five periods were used since these are the only periods in which there were 10 active bidders in every session (for the first iteration). For the second iteration we have 10 periods available (since there were 10 periods in the second iteration of the English clock auction). To make the two iterations more directly comparable, we report our regression results for the first five period of the first iteration (Table 3) and for the first five periods of the second iteration (Table 4), in which case, the same parameter values are used for each iteration. The regression results from using 10 periods in the second iteration are similar to those from using 5 periods.

In (1), the variable $\overline{v}_{pH}$ = 0 for periods in low CRT sessions and for high CRT sessions it is the difference between the maximum valuation across all bidders in period $p$ and the average of the maximum valuations across all periods $p \in \{1,2,3,4,5\}$. The variable $\overline{v}_{pL}$ = 0 for periods in high CRT sessions and for low CRT sessions it is the difference between the maximum valuation across all bidders in period $p$ and the average of the maximum valuations across all periods $p \in \{1,2,3,4,5\}$. The variable $x_{Cps} = 0.5$ for period $p$ if session $s$ is a high CRT session and $x_{Cps} = -0.5$ for period $p$ if $s$ is a low CRT session. The variable $x_{Tps} = 0.5$ for period $p$ if $s$ is a transparent (English clock) session and $x_{Tps} = -0.5$ for period $p$ if $s$ is a non-transparent (second price) session. Variable $x_{CTps} = 0.5$ for periods in transparent high CRT sessions and nontransparent low CRT sessions and $x_{CTps} = -0.5$ for periods in nontransparent high CRT sessions and transparent low CRT sessions. This approach parametrizes the effects of CRT and transparency as deviations from a mean effect size. The $x_{ps}$ variables, with $s \in \{1,2,3,4,5,6,7,8,9,10,11,12\}$ are session dummy variables that equal 1 if period $p$ is in session $s$ and 0 otherwise.

Regression model (1) enables us to identify the effects of cognitive reflection and mechanism transparency as well as their interaction in the context of bidding in auctions. The regression was implemented as a constrained regression in STATA 14.2 with the constraints that $\sum_s \kappa_s = 0$ for each treatment (with three sessions per treatment). Under this constraint, each third $\kappa_s$ value is determined given the values of the other two and so $\kappa_s$ values for sessions 3, 6, 9, and 12 do not appear in Tables 3 and 4.

Note that the standard economic prediction is that the valuation of the winning bidder should be fully determined by (and equal to) the maximum valuation among all bidders in the auction. Somewhat surprisingly, Table 3 reveals that, at least for the first five periods of Iteration 1, the maximum possible valuation is not related to the actual valuation of the winning bidder for either CRT group (p = 0.331 for high CRT subjects and p = 0.443 for low CRT subjects). That is, the only factor that should matter has no predictive power in the first iteration. However, both the CRT variable and the transparency variable are highly significant (as is their interaction). This suggests that in novel environments, variation in objective rationality as measured by the CRT is more predictive of market outcomes than is variation in economic fundamental values. In addition, for novel environments, mechanism transparency is more predictive of market outcomes than fundamental economic values.

From comparing Tables 3 and 4 we can also glean some insight into the effect of experience. For instance, when running model (1) on the first five periods of Iteration 2, CRT and transparency continue to be significant, while, the coefficients on $ln(\overline{v}_{pH})$ and $ln(\overline{v}_{pL})$ that were not significant in Iteration 1 are now both significant for Iteration 2. Thus maximum valuations do predict market outcomes in Iteration 2. Also, not that while CRT and transparency are significant for both iterations, the effect sizes for CRT and Transparency both decline considerably between iterations (a decline from 0.488 to 0.162 for the CRT effect size and from 0.578 to 0.166 for the transparency effect size), suggesting that effects of CRT and transparency are weaker when participants have experience with their environment.

**Table 3. Regression Results for Model (1): Auction Iteration 1**

| Meaning of estimated parameters | Estimated Values | Std. Error | p-value |
|---|---|---|---|
| | | | |
| Mean Effect Size across all treatments ($\alpha$) | 2.926 | 0.083 | < 0.001 |
| Effect of Valuation for High CRT ($\beta_H$) | 1.355 | 1.379 | 0.331 |
| Effect of Valuation for Low CRT ($\beta_L$) | 1.067 | 1.379 | 0.443 |
| Cognitive Reflection Parameter ($\gamma$) | 0.488 | 0.166 | 0.005 |
| Transparency Parameter ($\lambda$) | 0.578 | 0.166 | 0.001 |
| Interaction Effect ($\phi$) | -0.441 | 0.166 | 0.010 |

**Table 4. Regression Results for Model (1): Auction Iteration 2**

| Meaning of estimated parameters | Estimated Values | Std. Error | p-value |
|---|---|---|---|
| | | | |
| Mean Effect Size across all treatments ($\alpha$) | 3.158 | 0.032 | < 0.001 |
| Effect of Valuation for High CRT ($\beta_H$) | 1.107 | 0.539 | 0.045 |
| Effect of Valuation for Low CRT ($\beta_L$) | 1.634 | 0.539 | 0.004 |
| Cognitive Reflection Parameter ($\gamma$) | 0.162 | 0.065 | 0.005 |
| Transparency Parameter ($\lambda$) | 0.166 | 0.065 | 0.001 |
| Interaction Effect ($\phi$) | -0.171 | 0.065 | 0.010 |

## 4. Games with Complete Information

In addition to studying the effects of cognitive reflection, experience, and mechanism transparency on decision quality in games of incomplete information (Section 3), we also conducted experiments to study these effects in games of complete information and in individual choices under risk.

### 4.1 Experimental Design for Static and Dynamic Random Serial Dictatorships

As a game of complete information, following Li (2017) we employed two variants of a random serial dictatorship (RSD). In the RSD mechanism, agents know the payoffs available to other agents, they are assigned a random priority, and they receive their most preferred payoff available to them when there priority number is reached. The RSD is both strategyproof and efficient. We employ the two variants of an RSD mechanism due to Li (2017): (i) a dynamic RSD in which subjects are assigned a priority number and then take turns choosing prizes without replacement from a list of possible prizes (an OSP mechanism), and (ii) a static RSD in which subjects submit a ranking over all possible prizes and then receive the highest ranked prize on their list when their priority number is reached (an SP mechanism).

Our experimental design for the random serial dictatorships is partially a replication of the RSD experiment in Li (2017) but with the added dimension of studying the role of cognitive reflection and how it relates to experience and mechanism transparency. We administered twelve experimental sessions, with each session involving two iterations of ten rounds of an RSD mechanism, and two iterations of ten lottery choices (described in Section 5). As in the auction studies, we recruited high CRT sessions and low CRT sessions. In particular, we conducted six high CRT sessions: four sessions with nontransparent (static) RSD choices and nontransparent lottery choices and two sessions with transparent (dynamic) RSD choices and transparent lottery choices[7]. We also conducted six analogous low CRT sessions. Each session consisted of twelve subjects. The software and instructions for the RSD games were the same as those used in Li (2017). The instructions are also included in Appendix A.

In the RSD experiments, subjects could earn prizes from the set {$0, $0.25, $0.50, $0.75, $1.00, $1.25} in any given period. Four of these six prizes were randomly selected by the software to be allocated among subjects in each RSD game in each period. Each group of twelve subjects was randomly assigned into three groups of four subjects who submitted a ranking over the prizes (in the static RSD), or who took turns picking a prize (in the dynamic RSD). Each group persisted throughout the experiment so that the groups are the units of independent observations.

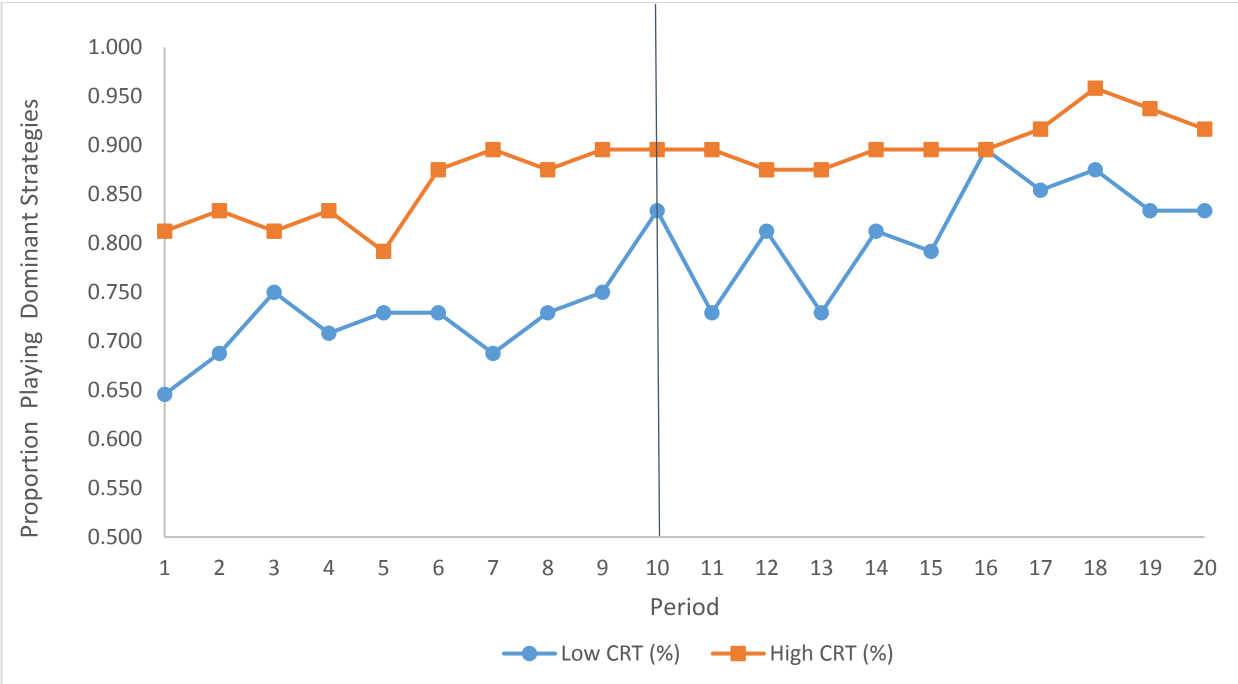
### 4.2 Experimental Results

The dynamic RSD mechanism involves rather trivial decisions, such as a choice between a bigger or a smaller amount of money. This implementation of an RSD mechanism makes the optimal strategy of choosing the biggest payoff transparent. While the static RSD mechanism is also very simple – and indeed, our data suggests most subjects found it to be so, it is decidedly less transparent than the dynamic RSD. In the static RSD, any possible ranking of prizes is permitted. If subjects realize they are in a potentially strategic situation, they would need to reason themselves to the dominant strategy of ranking payoffs according to their value. While such reasoning may typically be very natural, it is nevertheless an extra step or two of thinking beyond what is required for the dynamic RSD. In the dynamic RSD, one need only consider the available subset of the six prizes used in the experiment and pick the largest one.

Figure 6 displays the proportion of subjects playing the dominant strategy (of ranking payoffs from highest to lowest) across all 20 periods in the static RSD mechanism for both high CRT and low CRT participants. A baseline of 81.3% of the 48 High CRT subjects played the dominant strategy in Period 1, which increased to over 90% by Period 20. A baseline of 64.6% of the 48 Low CRT subjects played the dominant strategy in Period 1, which increased to nearly 85% by Period 20. For the first period, the difference in performance between high and low CRT participants is marginally significant ($p = 0.066$, two-tailed Z-difference in proportions test). High CRT subjects outperform low CRT subjects within each period. However, both High and Low CRT subjects converge toward equilibrium play with experience.

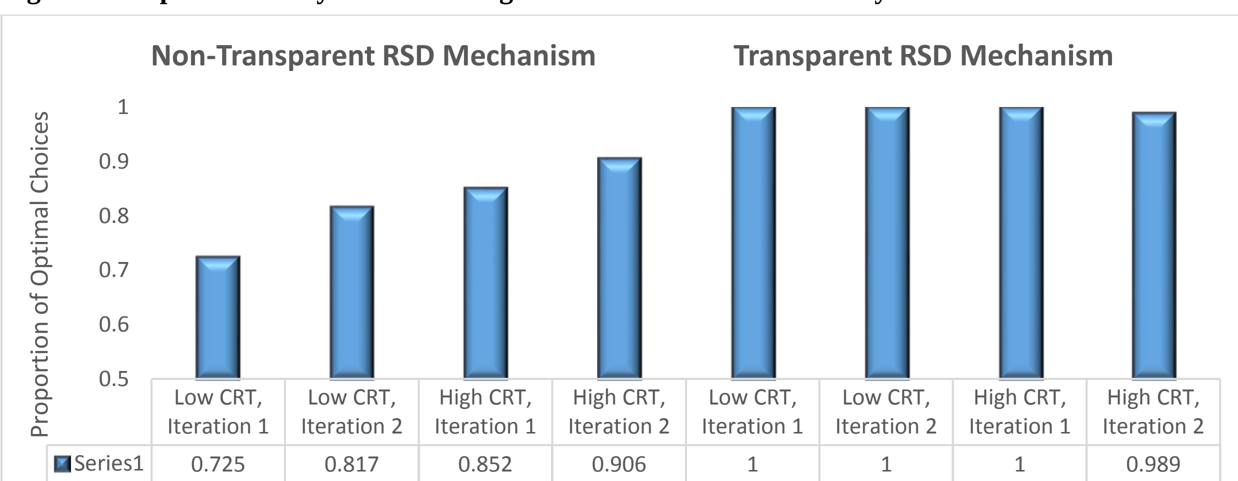
[7] We used a balanced design: In two of the four static RSD sessions and one of the two dynamic RSD sessions, the RSD iterations were conducted first. In the other three sessions, the lottery task was conducted first. The transparent RSD and lottery sessions involved trivial choices (e.g., choosing directly between $0.75 and $1.00), so we felt it unnecessary to conduct many of these sessions.

**Figure 6: Proportion Playing Dominant Strategies over Time in Static Random Serial Dictatorship**




Performance on the dynamic RSD mechanism (proportion of payoff-maximizing choices) is summarized in Figure 7 alongside the proportion of payoff-maximizing rankings for the static RSD mechanism. This data is provided for both high and low CRT subjects and for both the first iteration and second iteration of the mechanism (with 10 periods per iteration). From the figure, we see again that low CRT subjects with experience (81.7% payoff maximizing rankings) perform similarly to high CRT subjects without experience (85.2% payoff-maximizing rankings) for the less transparent (static) RSD mechanism. In addition, the more transparent mechanism (dynamic RSD) eliminates the gap in performance between high and low CRT subjects, even in the first iteration.

**Figure 7. Proportion of Payoff Maximizing Decisions for the Static and Dynamic RSD Mechanism**



### 4.3 Regression Model

To further analyze the RSD data, we implemented a simple regression model:

$$p_g = \alpha + \gamma x_{Cg} + \lambda x_{Tg} + \phi x_{CTg} + \varepsilon_g \tag{2}$$

where $p_g$ is the proportion of periods that achieved the dominant strategy equilibrium for group $g$. Model (2) also has a mean effect size, $\alpha$, a cognitive reflection parameter, $\gamma$, a transparency parameter $\lambda$, and a coefficient,$\phi$, to allow for an interaction effect between cognitive reflection and transparency. The variable $x_{Cg} = 0.5$ for high CRT groups and $x_{Cg} = -0.5$ for low CRT groups. The variable $x_{Tg} = 0.5$ for groups in transparent RSD sessions and $x_{Tg} = -0.5$ for groups in non-transparent RSD sessions. The variable $x_{CTg} = 0.5$ for groups in transparent high CRT sessions and nontransparent low CRT sessions and $x_{CTg} = -0.5$ for groups in nontransparent high CRT sessions and transparent low CRT sessions. As before, this approach parametrizes the effects of CRT and of transparency (and of their interaction) as deviations from a mean effect size.

The regression results are summarized in Table 5 (for Iteration 1) and in Table 6 (for Iteration 2). Note that none of the predicted values exceed one since the CRT, transparency, and interaction values each have a factor of 0.50 incorporated to them. From Tables 5 and 6, we see that CRT is not significant in either iteration. However, CRT does appear to have a systematic effect in that high CRT subjects performed better than low CRT subjects in the static RSD mechanism in every period of both iterations (see Figure 6). Moreover, in the first iteration (first 10 periods) of the static RSD mechanism, low CRT groups achieved the dominant strategy equilibrium in 34.2% of cases, whereas high CRT groups achieved the equilibrium in 57.5% of cases. Since the group is the unit of independent observations, having only 12 high CRT groups and 12 low CRT groups in the non-transparent RSD treatments might explain why this difference is not statistically significant.

Tables 5 and 6 also reveal that transparency is highly significant in both iterations. The transparent RSD mechanism performs significantly better than the non-transparent mechanism as predicted by Li (2017). Finally, experience also has an effect as the coefficient on both cognitive reflection and on transparency decrease from the first iteration to the second iteration.

**Table 5. Regression Results for Model (2): RSD Iteration 1**

| Meaning of estimated parameters | Estimated Values | Std. Error | p-value |
|---|---|---|---|
| | | | |
| Mean Effect Size across all treatments ($\alpha$) | 0.729 | 0.044 | < 0.001 |
| Cognitive Reflection Parameter ($\gamma$) | 0.117 | 0.088 | 0.195 |
| Transparency Parameter ($\lambda$) | 0.542 | 0.088 | < 0.001 |
| Interaction Effect ($\phi$) | -0.117 | 0.088 | 0.195 |

**Table 6. Regression Results for Model (2): RSD Iteration 2**

| Meaning of estimated parameters | Estimated Values | Std. Error | p-value |
|---|---|---|---|
| | | | |
| Mean Effect Size across all treatments ($\alpha$) | 0.790 | 0.049 | < 0.001 |
| Cognitive Reflection Parameter ($\gamma$) | 0.088 | 0.098 | 0.381 |
| Transparency Parameter ($\lambda$) | 0.388 | 0.098 | < 0.001 |
| Interaction Effect ($\phi$) | -0.121 | 0.098 | 0.229 |

## 5. Individual Choices under Risk

To further study the relationship between cognitive reflection, experience, and choice architecture, we employed a task involving individual choices between lotteries. In transparent sessions, subjects participated in the dynamic RSD mechanism and transparent lottery choices. In non-transparent sessions, subjects participated in the static RSD mechanism and non-transparent lottery choices.

### 5.1 Experimental Design

A total of 144 undergraduate students at a private California university participated in either a transparent or a non-transparent session. There were six high CRT sessions and six low CRT sessions, each set consisting of four non-transparent and two transparent sessions with 12 subjects per session.

In the lottery task in each session, subjects made individual decisions between 20 pairs of lotteries (two iterations of a fixed set of 10 distinct lottery pairs). The same 10 lottery pairs in the same sequence were used for all subjects in both the first and second iteration of both the transparent and non-transparent lottery tasks. The lottery pairs only differed in how they were framed. An example lottery pair is given in Figure 8 in both the nontransparent and the transparent choice architectures. The other lottery pairs in the experiment were similar and are provided in Appendix B. In each lottery pair, one lottery stochastically dominates the other. These pairs are similar to a lottery pair used by Tversky and Kahneman (1986), although to our knowledge, they have not been used in conjunction with the CRT or in a series of decisions where subjects can modify their behavior with experience. In the transparent lottery choices, the salient comparison favored the optimal lottery. In the non-transparent lottery choices, the salient comparison favored the stochastically dominated lottery.

The transparent sessions included the transparent RSD mechanism and transparent lottery frames. The non-transparent session included the non-transparent RSD mechanism and non-transparent lottery frames. The order of tasks (RSD versus lottery) was counterbalanced for both transparent and non-transparent sessions. All lottery choices were variants of those in Figure 8 (including those in the figure), with nontransparent sessions employing the representation in the top panel of Figure 8 and transparent sessions employing the representation in the bottom panel. Each lottery had four possible outcomes which ranged, across the ten lotteries, between \$0 and \$0.75. Subjects received feedback after each choice as the previous lottery pairs, a subject's previous choices, and the payoffs that subject received were displayed on that subject's screen below the current pair that the subject was choosing between. Once all subjects had recorded a choice, a spinner with 100 tick-marks would determine the 'winning number' for that choice. For instance, in the choice in Figure 8, any spinner number between 0 and 90 yielded a prize of \$0.30, for both the red and the blue option. Subjects were paid for each of their lottery choices.

The software randomized the color (red or blue) and position (top or bottom) of the lotteries. The sequence of lottery choices was fixed and the spinner numbers were pre-drawn so that all subjects received the same feedback in the same order. Moreover, the random draw of numbers produced the same payouts for both lotteries, for each of the ten lottery pairs. This was likely due to the close similarity between each lottery within a pair, with prizes only differing by one winning number in each pair. However, in each pair, one lottery first order stochastically dominated the other.

In a sense, the lottery experiment described here provides a stringent test of the hypothesis that people optimize. It is often argued that when high stakes, competition, and learning are present, rational behavior is more likely to emerge. In the present task, the lotteries involved small stakes (with no prize greater than \$0.75), they faced no competition, and there was no opportunity to learn from feedback (since feedback was the same regardless of which choice they made). It might then be surprising if people do learn to optimize, particularly, in the nontransparent task, given the absence of strong incentives, competition, and helpful feedback. Prior to making their lottery choices, subjects read the instructions at their own pace and responded to four quiz questions (each worth \$0.50), to check their understanding.

| Select One | Payout A | Winning #s | Payout B | Winning #s | Payout C | Winning #s | Payout D | Winning #s |
|---|---|---|---|---|---|---|---|---|
| ○ Red | 30¢ | 1 to 90 | 25¢ | 91 to 95 | 20¢ | 96 | 0¢ | 97 to 100 |
| ○ Blue | 30¢ | 1 to 90 | 25¢ | 91 to 96 | 0¢ | 97 | 0¢ | 98 to 100 |

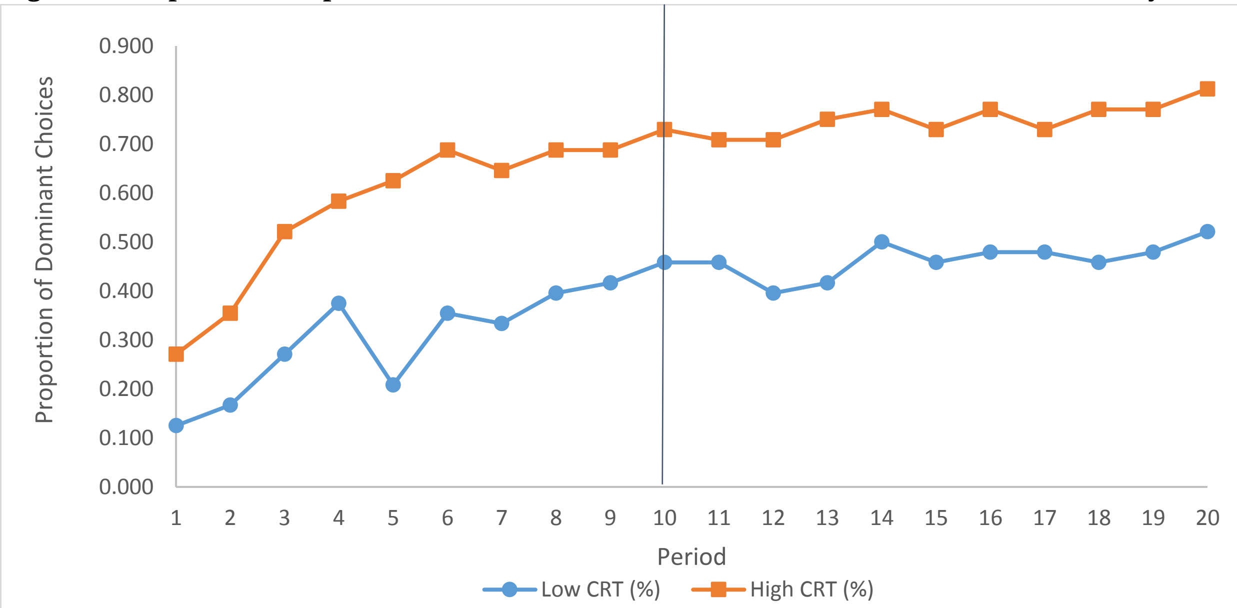
| Select One | Payout A | Winning #s | Payout B | Winning #s | Payout C | Winning #s | Payout D | Winning #s |
|---|---|---|---|---|---|---|---|---|
| ○ Red | 30¢ | 1 to 90 | 25¢ | 91 to 95 | 20¢ | 96 | 0¢ | 97 to 100 |
| ○ Blue | 30¢ | 1 to 90 | 25¢ | 91 to 95 | 25¢ | 96 | 0¢ | 97 to 100 |

**Figure 8. Nontransparent Choice Architecture (Top) and Transparent Architecture (Bottom).** The salient comparison in the nontransparent architecture (20 vs. 0 cents) favors the dominated lottery.

### 5.2 Experimental Results

The proportion of optimal choices made by high CRT and low CRT subjects across all 20 periods is displayed in Figure 9 for the non-transparent lottery task (with a vertical line separating the periods from the first and second iterations). Note that in every period of both iterations in Figure 9, high CRT subjects performed better than low CRT subjects. There were 12.5% of 48 low CRT subjects and 27.1% of 48 high CRT subjects who chose the optimal (stochastically dominant) lottery in Period 1. This difference is marginally significant (two-tailed Z difference in proportions test, $p = 0.073$).

**Figure 9: Proportion of Optimal Choices over Time when Salience favors Dominated Lottery**

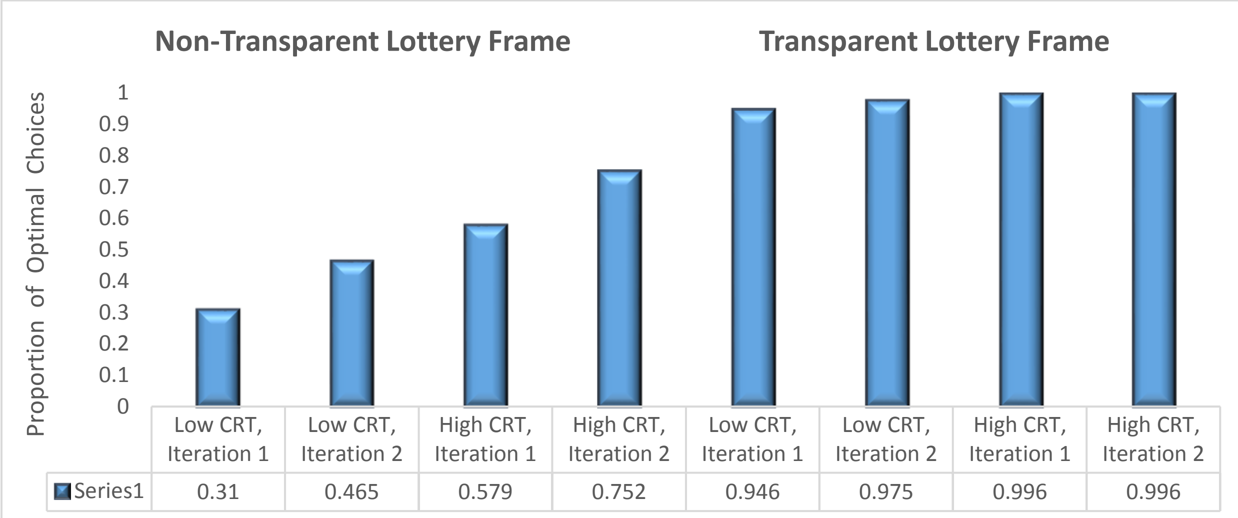



For both CRT groups, a large majority of subjects in our experiment chose the dominated option on their first choice in the non-transparent task. Two reasons for this choice are that (i) one's attention is naturally drawn to the most salient difference between lotteries which is the difference between 0 and 20 cents that favors the dominated lottery in the top panel of Figure 8, and (ii) the dominated lottery (red) has more outcomes displayed which pay more than 0. However, upon inspection it is clear that the blue option has the dominating probability distribution as it offers at least as good a prize as red at every probability level and offers a strictly better prize at some probabilities.

The lottery task is not an environment where feedback is likely to help modify behavior. The difference in expected values between the lotteries is small and one could conjecture that the cognitive costs of computing the values of such complex lotteries is not worth the small expected gain. One might contrast this with the second price auction in which strong feedback (going bankrupt) and strong incentives (larger monetary payoffs) naturally push behavior towards the dominant strategy. Conditions which are thought to induce optimizing behavior – significant monetary incentives, competition, and experience are all present in the second price auction. Of these conditions, only experience is present for the lottery task. Given the preceding comments, we find it remarkable that selection of the optimal lottery increased from 12.5% to 52.1% for low CRT subjects and from 27.1% to 81.3% for high CRT subjects.

Summary metrics for the lottery task are also provided in Figure 10. We again see that low CRT subjects with experience are not very different from high CRT subjects without experience in the nontransparent task. However, this gap is a little larger than the gaps for the auctions and the RSD tasks. Also, as before, transparency reduces the gap in performance between high and low CRT subjects.

**Figure 10: Proportion of Optimal Lottery Choices Across Treatments and Iterations**




## 5.3 Regression Model

To further analyze the lottery choice data, we implemented a simple regression model:

(3) $$p_s = \alpha + \gamma x_{Cs} + \lambda x_{Ts} + \phi x_{CTs} + \varepsilon_s$$

where $p_s$ is the proportion of periods in which subject $s, s \in \{1,2,\ldots,144\}$ chose the dominant lottery. Model (3) also has a mean effect size, $\alpha$, a cognitive reflection parameter, $\gamma$, a transparency parameter $\lambda$, and a coefficient,$\phi$, to allow for an interaction effect between cognitive reflection and transparency. The variable $x_{Cs} = 0.5$ for high CRT subjects and $x_{Cs} = -0.5$ for low CRT subjects. The variable $x_{Ts} = 0.5$ for subjects in transparent sessions and $x_{Ts} = -0.5$ for subjects in non-transparent sessions. The variable $x_{CTs} = 0.5$ for subjects in transparent high CRT sessions and nontransparent low CRT sessions and $x_{CTs} = -0.5$ for subjects in nontransparent high CRT sessions and transparent low CRT sessions. As before, this approach parametrizes the effects of CRT and of transparency (and of their interaction) as deviations from a mean effect size.

The regression results are summarized in Table 7 (for Iteration 1) and in Table 8 (for Iteration 2). Note that none of the predicted values exceed one since the CRT, transparency, and interaction values each have a factor of 0.50 incorporated to them. From Tables 7 and 8, we see that CRT is significant in both iterations, with similar parameter estimates. Tables 7 and 8 also reveal that transparency is significant in both iterations. Finally, experience also has an effect as the coefficient on both cognitive reflection and on transparency decrease from the first iteration to the second iteration. This decrease is very small for cognitive reflection but quite large for transparency. Moreover, across each of the regression models (1), (2), and (3), switching from non-transparent to transparent mechanisms has a greater effect on performance than switching from low cognitive reflection subjects to high cognitive reflection subjects.

**Table 7: Regression Results for Model (3): Lottery Iteration 1**

| Meaning of estimated parameters | Estimated Values | Std. Error | p-value |
|---|---|---|---|
| | | | |
| Mean Effect Size across all treatments ($\alpha$) | 0.708 | 0.025 | < 0.001 |
| Cognitive Reflection Parameter ($\gamma$) | 0.159 | 0.051 | 0.002 |
| Transparency Parameter ($\lambda$) | 0.526 | 0.051 | < 0.001 |
| Interaction Effect ($\phi$) | -0.109 | 0.051 | 0.033 |

**Table 8: Regression Results for Model (3): Lottery Iteration 2**

| Meaning of estimated parameters | Estimated Values | Std. Error | p-value |
|---|---|---|---|
| | | | |
| Mean Effect Size across all treatments ($\alpha$) | 0.800 | 0.031 | < 0.001 |
| Cognitive Reflection Parameter ($\gamma$) | 0.154 | 0.062 | 0.015 |
| Transparency Parameter ($\lambda$) | 0.377 | 0.062 | < 0.001 |
| Interaction Effect ($\phi$) | -0.133 | 0.062 | 0.035 |

## 6. Discussion

We investigated how performance in strategyproof mechanisms and individual lottery choices is affected by choice architecture, experience, and cognitive reflection. We found that experience serves as a substitute for cognitive reflection: For each of the mechanisms in this experiment, low CRT subjects with experience performed similarly to high CRT subjects without experience. Further, for novel environments, we find the CRT to be reliable in sorting out heterogeneity in objectively rational decisions across individual lottery choices, random priority mechanisms, and auctions. Differences in cognitive reflection, as measured by the CRT, provide a unified explanation for differences in objectively rational behavior in individual and strategic decisions. Given this relationship between cognitive reflection and objective rationality, a promising area for research is to model agents who vary in objective rationality, analogous to standard approaches for modeling heterogeneity in subjective rationality (risk preferences, time preferences, and beliefs). Zhang and Levin (2017) provide a promising step in this direction.

From a practical perspective, the CRT is important because the types of errors it measures may arise in the real world. For instance, in the lottery task the intuitive response seems to involve focusing on the salient payoff difference and choosing the alternative with the larger salient payoff. Some reflection is needed to detect the dominance relation. Such decisions may share similarities with consumer purchase decisions where looking casually or carefully at sale prices and product advertisements may determine whether a consumer avoids purchasing a dominated option that is cleverly ‘framed’ in the retailer’s ad.

Another question to ask is: When does cognitive reflection matter? We conclude that cognitive reflection matters most in novel environments (i.e., in which agents lack experience) and in complex environments (in which there are many options or outcomes). We observed that gaining familiarity with the choice environment (through experience) or simplifying the choice environment (through choice architecture) leads behavior to converge toward the normative standard of objective rationality.

## References


Branas-Garza, P., Garcia-Munoz, T., and Hernan-Gonzalez, R. (2012). Cognitive effort in the beauty contest game. *Journal of Economic Behavior and Organization*, **83**, 254-260.

Burnham, T.C., Cesarini, D., Johannesson, M., Lichtenstein, P., and Wallace, B. (2009). High cognitive ability is associated with lower entries in a p-beauty contest. *Journal of Economic Behavior and Organization*, **72**, 171-175.

Corgnet, B., M. DeSantis, D. Porter. (2015). Revisiting Information Aggregation in Asset Markets: Reflective Learning & Market Efficiency. ESI Working Paper 15-15

Cox, J., Roberson, R., and Smith, V.L. (1982). Theory and behavior of single object auctions. In *Research in Experimental Economics*, vol. 2, ed. Smith, V.L. (JAI Press, Greenwich, CT), pp. 1-43.

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

Gilboa, I., Maccheroni, F., Marinacci, M., and Schmeidler, D. (2010). Objective ad subjective rationality in a multiple prior model. *Econometrica*, **78**, 755-770.

Kagel, J.H. and Levin, D. (1993). Independent private value auctions: Bidder behavior in first-, second-, and third-price auctions with varying numbers of bidders. *The Economic Journal*, **103**, 868-879.

Kagel, J.H., Harstad, R.M., and Levin, D. (1987). Information impact and allocation rules in auctions with affiliated private values: a laboratory study. *Econometrica*, **55**, 1275-1304.

Kagel, J.H. (1995). Auctions: a survey of experimental research. In *The Handbook of Experimental Economics,* eds. Kagel, J.H. & Roth, A.E. Princeton, 501-585.

Levitt, S. D., List, J. A., and Sally, E. (2011). Checkmate: Exploring backward induction among chess players. *The American Economic Review*, **101**, 975-990.

Li, S. (2017). Obviously strategy-proof mechanisms. *The American Economic Review*, **107**, 3257-3287.

Schnusenberg, O. and Gallo, A. (2011). On cognitive ability and learning in a beauty contest. *Journal for Economic Educators*, **11**, 13-24.

Toplak, M.E., West, R.F., and Stanovich, K.E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, **20**, 147-168.

Tversky, A. and Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, **59**, S251-S278.

Vickrey, W. (1961). Counter speculation, auction and competitive sealed tenders. *Journal of Finance*, **16**, 8-37.

Zhang, L., and Levin, D. (2017). Partition Obvious Preference and Mechanism Design: Theory and Experiment. Manuscript.