Adverse and Advantageous Selection in the Laboratory*

S. Nageeb Ali† Maximilian Mihm‡ Lucas Siga§ Chloe Tergiman¶

November 20, 2020

Abstract

We study two-player games where one-sided asymmetric information can lead to either adverse or advantageous selection. We contrast behavior in these games with settings where both players are uninformed. We find stark differences, suggesting that subjects do account for endogenous selection effects. Removing strategic uncertainty increases the fraction of subjects who account for selection. Subjects respond more to adverse than advantageous selection. Using additional treatments where we vary payoff feedback, we connect this difference to learning. We also observe a significant fraction of subjects who appear to understand selection effects but do not apply that knowledge.

JEL: C91, C92, D72, D81, D82.

Keywords: Laboratory experiment, adverse selection, advantageous selection, contingent reasoning, strategic uncertainty, voting, social preferences.

*We thank the co-editor and three anonymous referees for their thoughtful and constructive feedback. We thank Colin Camerer, Juan Carrillo, Alex Imas, Ryan Oprea, Emanuel Vespa, and Sevgi Yuksel for useful comments and suggestions. Robizon Khubulashvili provided excellent research assistance and programming. We are grateful to REF grant #RE099 for funding this experiment.

†Department of Economics, Penn State. Email: nageeb@psu.edu.

‡Division of Social Science, NYU-Abu Dhabi. Email: max.mihm@nyu.edu.

§Division of Social Science, NYU-Abu Dhabi. Email: lucas.siga@nyu.edu.

¶Smeal College of Business, Penn State. Email: cjt16@psu.edu.

Contents

1 Introduction

Motivation: Asymmetric information is central to many economic and social interactions. When individuals are asymmetrically informed, it can be rational for the less informed individual to be suspicious of the motives of someone who is better-informed. For instance, Akerlof (1970) illustrates how buyers should be pessimistic about the quality of products being sold given that better informed sellers are willing to sell those objects. Rothschild and Stiglitz (1976) argue that insurance providers should set premiums anticipating that those who privately know that they have a higher likelihood of claiming the insurance also have a greater incentive to buy it. Similarly, the "No-Trade Theorem" (Milgrom and Stokey, 1982) articulates how bettors engaged in speculative trading should draw inferences based on the motive that others have for taking opposing bets.

Across these settings, we see a common theme of *adverse selection*. From the perspective of each individual, the payoff of an available option—be it buying used cars, selling insurance, or taking a bet—is determined both by nature and the endogenously chosen actions of other parties. But selection is not always adverse; when preferences are aligned, then the selection may be *advantageous*. For instance, if potential insurees are better-informed about their risk preferences, those who have a higher demand for insurance may be the risk-averse individuals who are "good risks" for an insurer (De Meza and Webb, 2001; Fang, Keane, and Silverman, 2008). In elections where voters share common preferences, some voters may be willing to abstain on ballot propositions to let better-informed voters cast the decisive votes, and thereby benefit from the selection of outcomes generated by the actions of others (Feddersen and Pesendorfer, 1996).

We study how well people account for adverse and advantageous selection. How do people behave when they know that they are asymmetrically informed? Is the impact of asymmetric information uniform across adverse and advantageous selection or do people account for selection more in some settings than in others?

Main Design and Findings: The core of our design is a simple two-player game described in Section 2: Alice and Bob jointly choose between a safe and risky option. The safe option yields identical payoffs for each party. The risky option is a lottery with a high payoff that exceeds that of the safe option and a low payoff below it. In *positively correlated* rounds, Alice and Bob obtain identical payoffs from the risky option and so either both gain or lose from the risky option being chosen. Preferences here are perfectly aligned. In *negatively correlated* rounds, Alice and Bob have misaligned interests: relative to the safe option, one of them gains from the risky option while the other loses. Ex ante, each is equally likely to be the winner. In both positively and negatively correlated rounds, players vote simultaneously between the safe and risky option, and the risky option is selected if and only if both players vote for it. Importantly, one player—say Alice—privately observes the realized payoffs of the risky option whereas the other player is only told whether it is positively or negatively correlated. That the pair is asymmetrically informed is common knowledge between them.

What do standard theories of selection predict in this setting? If players are selfish and play weakly undominated strategies, the informed player (Alice) votes for the risky option if and only if it benefits her. Anticipating this choice, Bob always votes for the risky option if payoffs are positively correlated because the risky option is then selected advantageously: if he votes for it, then it is selected only when Alice votes for it as well, which means that it must benefit her and therefore him too. By contrast, if payoffs are negatively correlated, then Bob always votes for the safe option because he anticipates that the risky option is selected *adversely*: Alice votes for the risky option only when she gains from it, which means he must lose from it. These behavioral predictions for the two asymmetric information games do not require the fixed-point logic of equilibrium but instead follow from two rounds of elimination of weakly dominated strategies.

This reasoning suggests a test of whether subjects account for selection: when they are in the role of the uninformed player (Bob), are they more likely to choose the risky option when payoffs are positively correlated than when they are negatively correlated? We show in Table 3 that the answer is yes. We consider two payoff variations for each correlation condition: one where the safe option yields a low payoff (below the expected value of the risky option) and one where it yields a high payoff (above the expected value of the risky option).[1] When the safe option has a low value, shifting the payoffs from negative to positive correlation raises the fraction choosing the risky option from 48% to 86%; when the safe option has a high value, the corresponding numbers are 1% and 33%. Thus, behavior shifts in the direction predicted by theories of asymmetric information. Moreover, about 20% of subjects make every choice in a way that is fully consistent with the predictions of adverse and advantageous selection.

We control for non-informational confounds by comparing the above behavior with games where both players learn the correlation structure but are symmetrically uninformed about the realized payoffs.[2] In this game, shifting from negative to positive correlation

[1] We randomize the order across subjects.

[2] One potential confound is aversion to inequality: the negatively-correlated risky option is ex ante fair but ex post unequal whereas the positively-correlated risky option is both ex ante and ex post equal. If subjects are averse to ex post inequality, there is a confounding rationale for a subject to choose the safe option when payoffs are negatively correlated but not when payoffs are positively correlated.

raises the fraction voting for the risky option from 78% to 88% with a low payoff for the safe option, and from 4% to 8% with a high payoff for the safe option. Thus, we see shifts in the same direction, but with a much smaller magnitude. Comparing the magnitude of these effects with those of the asymmetric information games suggests that a significant fraction of subjects do account for asymmetric information.

However, we also find that a significant proportion of subjects fail to account for selection, at least in some cases. Studying positive and negative correlation in a unified framework allows us to compare how subjects respond to adverse versus advantageous selection. We see that subjects respond more to adverse selection. When payoffs are negatively correlated, over half of the subjects consistently choose the safe option but when payoffs are positively correlated, less than a third of the subjects consistently choose the risky option. We investigate what accounts for this gap and, more generally, why behavior departs from the theoretical predictions of asymmetric information. To this end, we investigate the role of strategic uncertainty, difficulties of contingent reasoning, and a lack of payoff feedback about counterfactuals.

Strategic Uncertainty and Contingent Reasoning: Although behavior in our game is pinned down by eliminating two rounds of weakly dominated strategies when players are selfish, subjects may potentially face strategic uncertainty about the preferences and behavior of others. To assess the role of strategic uncertainty, we conduct a second treatment, described in Section 5, in which subjects are never paired with each other and are instead paired with computerized "robot" players whose strategies are known ahead of time. In the main asymmetric information game, these robot players observe the realized payoff and chooses the risky option if and only if it generates a higher (virtual) payoff for the robot than the safe option; human players never observe the realized payoff but know its correlation.

Removing strategic uncertainty significantly increases the degree to which subjects account for selection. Indeed, in this second treatment, when the safe option has a low value and subjects face a negatively correlated risky option, 77% of our subjects correctly choose the safe option, which is significantly higher than in the first treatment (52%). Similarly, when the safe option has a high value and subjects face a positively correlated risky option, the fraction of subjects who correctly choose the risky option is 46%, which is significantly higher than the proportion who do so in the first treatment (33%). The fraction of subjects who behave according to our asymmetric information predictions in all rounds almost doubles to 40%. Thus, strategic uncertainty captures (to a significant degree) a divergence between the selection effects we see in "human-human" interactions and those predicted by theory.

We also use this treatment to see if subjects have difficulties with the contingent reasoning required to determine selection effects. After subjects play against robots, they are asked several non-leading questions about the inferences they can draw from the robot's choice. These are relatively high stakes questions that deliver a high payment only if subjects answer every question correctly. The questions both measure how well subjects understand the relevant contingent-reasoning, and potentially provide subjects with a nudge that alters how they play the game. After answering these questions, the subjects play the asymmetric information game against the robot players once more.

Almost 90% of subjects answer all of the contingent-reasoning questions correctly. We find evidence that answering these contingent-reasoning questions increases the fraction of subjects who account for selection, but a significant fraction of subjects continue to deviate from theoretical predictions. Of the subjects whose choices depart from theoretical predictions the second time that they play the asymmetric information game, over three quarters answer all of the contingent reasoning questions correctly. These subjects appear able to understand each step of contingent reasoning separately but do not piece together that understanding in their subsequent strategic behavior.

Despite removing strategic uncertainty, we still see that subjects respond more to adverse selection than advantageous selection. In the Human-Robot treatment, 74% of subjects choose the safe option in every negatively-correlated round, but only 43% choose the risky option in every positively correlated round. This finding suggests a contextual aspect of contingent reasoning where people appear to account for contingencies in some settings but not in others. We show that this gap is not easily reconciled by models of limited strategic thinking, such as cursed equilibrium or level-k, and we conjecture that an alternative mechanism could contribute to the gap, which we turn to below.

Payoff Feedback: It is likely that people learn how to respond to asymmetric information based on experience. However, a challenge that people face is that they rarely observe counterfactuals: one observes the consequences only of actions that have been chosen, not of actions that have not been chosen. We view this inability to learn from counterfactuals as a potential source of asymmetry that can impact how subjects learn to respond to adverse and advantageous selection.

Here is why. If an individual consistently chooses a risky option in settings where payoffs are negatively correlated, she would repeatedly see that she is worse off than were she to choose the safe option. In everyday life, this is the mistake of "trusting" others when one shouldn't, and it is self-correcting in the long-run because one obtains the payoff feedback from that mistake. By contrast, if an individual consistently chooses a safe option when payoffs are positively correlated, then she does not observe what would have happened were she to have chosen the risky option instead. Experience simply does not teach her that this is a mistake. In everyday life, this is the mistake of not trusting others when one should, and this mistake is not self-correcting because one does not see the counterfactual.

We formalize this distinction using the language of self-confirming equilibrium (Fudenberg and Levine, 1993), which allows players to hold incorrect beliefs about the play of others so long as those beliefs are consistent with their payoff feedback. If payoffs are negatively correlated, then in every self-confirming equilibrium, the uninformed player must choose the safe option. The logic is that if the uninformed player were to choose risky, she obtains the payoff feedback that suggests that she is better off choosing safe. By contrast, if payoffs are positively correlated, then there exists a self-confirming equilibrium in which the uninformed player chooses the safe option. Upon choosing the safe option, the player does not obtain feedback that suggests it was the wrong choice. Thus, a failure to account for advantageous selection *can* be rationalized by incorrect beliefs off the equilibrium path while a failure to account for adverse selection *cannot*.

This reasoning suggests a natural test: if we vary whether subjects obtain information about off-path events, it should not affect behavior in negatively correlated rounds but should do so in positively correlated rounds. This is how our third and fourth treatments proceed. In the *partial feedback* treatment, in each round, after subjects make their decisions, they observe the payoff that they would obtain should that round be selected for payment. By contrast, in the *full feedback* treatment, subjects learn not only the information from the partial feedback treatment but also the realized payoff of the risky option and how the other player voted. Thus, even if subjects choose the safe option, they see the counterfactual outcome of what would have happened had they voted for the risky option. After these feedback rounds, subjects again play asymmetric information games without feedback. We see whether subsequent behavior is affected by the nature of previous feedback (partial or full).

We find that after partial feedback, 78% of subjects choose the safe option in every negatively-correlated round, and after full feedback, the corresponding proportion is 82%, a difference that is statistically insignificant. By contrast, if payoffs are positively correlated, 63% of subjects choose the risky option after partial feedback, and after full feedback, this proportion is 76%, which is a statistically significant difference. Moreover, there remains a significant gap, both statistically and in magnitude, in how subjects respond to adverse and advantageous selection after partial feedback (78% versus 63%, respectively) whereas this gap is statistically insignificant with full feedback (82% versus 76%, respectively). Thus, giving subjects feedback about counterfactuals reduces the gap between how well subjects account for adverse and advantageous selection.

We view this finding to be of both theoretical interest and germane to policy. Because, in practice, people do not observe counterfactuals, there may be a self-reinforcing cycle whereby individuals learn to distrust those who are better informed (from experiences when preferences are misaligned) and do not learn to rely on others when there are common gains. Our finding shows how *zero-sum thinking*—namely the tendency for people to treat strategic interactions as zero-sum games (Meegan, 2010; Różycka-Tran, Boski, and Wojciszke, 2015)—may persist and even be amplified by opportunities to learn. Based on their past experience in settings with asymmetric information, individuals may learn to correctly identify not to trust others when preferences are misaligned but not learn that they should behave differently in settings with common interests. This process suggests a direct consequence for political and electoral behavior. Given the widespread perception of polarization (Levendusky and Malhotra, 2015), relatively uninformed voters may believe that their interests are misaligned with those of better informed voters. Their suspicion may then lead them to vote in such a way that the election cannot be swung by the choices of better informed voters.[3]

Related Literature: A rich literature studies how people respond to asymmetric information in strategic settings including lemons markets (Bazerman and Samuelson, 1983), betting (Sonsino, Erev, and Gilat, 2002; Carrillo and Palfrey, 2011; Magnani and Oprea, 2017), settlements in zero-sum games (Carrillo and Palfrey, 2009), auctions (Kagel and Levin, 1986; Charness and Levin, 2009), elections (Guarnaschelli, McKelvey, and Palfrey, 2000; Battaglini, Morton, and Palfrey, 2010), and many others. Relative to the literature, we see the distinguishing features of our paper to be: (i) we compare behavior in asymmetric information games with otherwise identical games in which players are symmetrically informed to see whether subjects account for selection, (ii) we compare behavior in human-human interaction with human-robot interaction to investigate the role of strategic uncertainty in how subjects account for selection, and (iii) we investigate why subjects may or may not account for selection uniformly across settings with adverse and advantageous selection, and highlight the role of learning about counterfactuals.

One approach in the prior literature roots individual failures to account for asymmetric

[3]This behavior contrasts with that of Feddersen and Pesendorfer (1996) where uninformed voters abstain to let better informed voters swing the election. Ali, Mihm, and Siga (2018) show that negatively correlated payoffs can generally cause failures of information aggregation.

information to strategic uncertainty or incorrect beliefs about how others play the game. Brocas, Carrillo, Wang, and Camerer (2014) distinguish between these models in asymmetric information games by using "mousetracking" to record which payoffs subjects look at, and find support for theories where players are imperfectly attending to relevant information. We contribute to this perspective by seeing the degree to which subjects account for selection in both playing against human players as well as against robot players whose strategies are revealed ahead of time. We find that removing strategic uncertainty nearly doubles the fraction of subjects who account for selection. Yet, a significant fraction still fail to account for selection, and do so to a higher degree when there is advantageous selection. Our finding that payoff feedback matters in resolving the discrepancy between how much subjects account for adverse versus advantageous selection suggests that even when human subjects are told the strategies of robot players, experience is essential for them to "trust" the other player to make the right choice.

A recent literature studies failures in contingent-reasoning and selection-neglect; for example, see Esponda and Vespa (2014, 2018, 2019), Martínez-Marquina, Niederle, and Vespa (2019), Barron, Huck, and Jehiel (2019), and Enke (2019). Relative to this literature, we directly test whether people respond to asymmetric information by comparing choices where players are symmetrically uniformed with those where one player has private information. We see in this comparison that the behavior of an uninformed player changes when he knows that his opponent has private information, and this change is qualitatively in the direction predicted by theory albeit with a smaller magnitude. We also see that the degree to which people account for selection varies between adverse and advantageous selection, and our analysis suggests how payoff feedback influences the degree to which subjects account for selection.

We model the role of learning through self-confirming equilibria, where behavior is rationalized by potentially incorrect conjectures about off-path play. We vary whether subjects obtain feedback about off-path behavior and evaluate how such feedback affects subsequent behavior. We find support for self-confirming equilibria, complementing existing studies (Fudenberg and Levine, 1997; Fudenberg and Vespa, 2019).

2 A Conceptual Framework

This section describes the conceptual framework, which is also the central element of our design. Two players, Alice and Bob, simultaneously vote between two options, S (a safe option) and R (a risky option). The risky option R is selected if and only if both vote for it. The safe option S pays $s > 0$ to each of them. By contrast, R offers payoffs of l or h to each player where $0 < l < s < h$, and this lottery is implemented by the toss of a (virtual) fair coin. We denote a vector of payoffs by (π_A, π_B) where π_A is the amount paid to Alice and π_B is the amount paid to Bob. We vary whether R is positively or negatively correlated:

1. **Positive Correlation:** If the coin toss is *Heads*, R pays (l, l), and otherwise, R pays (h, h).
2. **Negative Correlation:** If the coin toss is *Heads*, R pays (l, h), and otherwise, R pays (h, l).

Positive correlation reflects a pure common-values environment in which every realization and choice guarantees that the players have equal payoffs. By contrast, in the negatively-correlated case, the risky option benefits one player to the detriment of the other (relative to the safe option).

In all of our experiments, subjects are told about the correlation of the risky option so they both know the possible payoffs of the risky option. Our setting of interest is one where information is asymmetric: Alice is told the realization of the coin toss, Bob is not, and this is common knowledge. In other words, Bob knows the *potential payoffs* (and the associated probability distribution) of the risky option whereas Alice knows the actual *realized payoffs* of the risky option.

Let us describe the strategic logic of this setting assuming that each player is selfish and has preferences represented by a utility function that is strictly increasing in wealth. We consider equilibria in weakly undominated strategies.[4] For both positively and negatively correlated payoffs, Alice has a unique weakly undominated strategy: vote for the risky option if she would obtain the high amount $h > s$ from it and for the safe option if she would obtain the low amount $l < s$ from the risky option. What does this imply for Bob? Assuming Alice plays this strategy, Bob's vote affects the outcome if and only if Alice is voting for the risky option because otherwise the safe option is selected regardless of his vote. So in the case where his vote matters, Alice must be obtaining a payoff of h if the risky option is selected. In the positive-correlation case, this is *advantageous selection* for Bob because he too must be obtaining the high amount h from the risky option, which makes voting for it a best response for him. By contrast, in the negative correlation case, this is *adverse selection* for Bob because then he must be obtaining l from the risky option, which makes voting for the safe option a best response for him. Thus, the

[4] There always exist equilibria in which both players choose S with probability 1 because the other is doing so. These equilibria are in weakly dominated strategies, and are not trembling-hand perfect.

equilibrium predictions are simple, and are pinned down by two iterations of eliminating weakly dominated strategies. We summarize below.

Proposition 1. *There exists a unique strategy profile that survives two rounds of elimination of weakly dominated strategies:*

(a) *The informed player* (Alice) *votes for the risky option if she obtains h from the risky option and votes for the safe option if she would obtain l from the risky option;*

(b) *The uninformed player* (Bob) *votes for the risky option if payoffs are positively correlated and for the safe option if payoffs are negatively correlated.*

This conceptual framework predicts that we should see the risky option being selected more often by an uninformed player in the positively-correlated case than in the negatively-correlated case. One may envision other rationales for this behavior (e.g., aversion to ex-post inequality), and our design disentangles the selection-motive from these other rationales.

3 Design and Procedures

This section describes our first treatment, namely the "Human-Human" (HH) treatment, where subjects were matched in pairs. Our second treatment—where subjects were instead matched with robot players—is described in Section 5.

3.1 Experimental Design

We described the *Asymmetric Information* (AI) game, in Section 2. We vary three elements of this game: (1) the payoff of the safe option S; (2) whether the risky option R is positively or negatively correlated; and (3) the identity of the player who learns the realized payoffs of the risky option R. The payoff of the safe option S, denoted by s in Section 2, is either \$12 or \$16 (for both parties). The values for l and h in the risky option R are \$10 and \$20, respectively, and the ex-ante probability that a subject receives either payoff if the risky option is implemented is set to 50%. Subjects played 8 rounds of this game, four where they were uninformed, and four where they perfectly learned the realized payoffs of R. These are summarized by Table 1.

Our objective is to assess the degree to which subjects account for selection effects. Following our theoretical predictions in Section 2, do subjects in the role of the uninformed player vote for S when it is negatively correlated and vote for R when it is positively

Table 1: Rounds in the Asymmetric Information game.

Round	Safe Option S (1 vote)	Risky Option R (2 votes)	Voter Informed	Other Voter Informed
12N	(\$12; \$12)	(\$10, \$20) or (\$20, \$10)	no	yes
12P	(\$12; \$12)	(\$10, \$10) or (\$20, \$20)	no	yes
16N	(\$16; \$16)	(\$10, \$20) or (\$20, \$10)	no	yes
16P	(\$16; \$16)	(\$10, \$10) or (\$20, \$20)	no	yes
12N	(\$12; \$12)	(\$10, \$20) or (\$20, \$10)	yes	no
12P	(\$12; \$12)	(\$10, \$10) or (\$20, \$20)	yes	no
16N	(\$16; \$16)	(\$10, \$20) or (\$20, \$10)	yes	no
16P	(\$16; \$16)	(\$10, \$10) or (\$20, \$20)	yes	no

correlated *because they are strategically accounting for selection*? To answer this question, we have to distinguish the asymmetric-information rationale for this behavior from other rationales for the same behavior. The other parts of the Human-Human treatment are designed with this goal in mind, allowing us to make within-subject comparisons across several games.

A confounding consideration is *aversion to ex post inequality*: the payoffs from R are ex post unequal when it is negatively correlated and ex post equal when it is positively correlated. By contrast, the payoffs from S are always ex post equal. To assess how much subjects are influenced by this consideration, we precede the AI game with the *Symmetric Information* (SI) game, which uses the same parameters as the AI game, but where players are symmetrically uninformed. That is, in the SI game, neither player is informed about the payoffs of option R, other than knowing its correlation structure. Because both players are symmetrically uninformed (and this is common knowledge), there is neither adverse nor advantageous selection in this game.

To evaluate the strength of social preference considerations (both aversion to ex post inequality and *preferences for efficiency*) without the interference of a voting structure, we had subjects play a series of Dictator games following the AI game. Table 2 shows the rounds that subjects faced in the Dictator games. Rounds 1 through 4 of the Dictator games directly correspond to the 12N, 12P, 16N and 16P rounds in both the AI and SI games. Rounds 5 through 8 allow us to evaluate subjects' preferences with respect to efficiency tradeoffs without the presence of uncertainty.[5]

Table 2: Rounds in the Dictator Game.

Round	Option A	Option B
1	(\$12; \$12)	(\$10, \$20) or (\$20, \$10)
2	(\$12; \$12)	(\$10, \$10) or (\$20, \$20)
3	(\$16; \$16)	(\$10, \$20) or (\$20, \$10)
4	(\$16; \$16)	(\$10, \$10) or (\$20, \$20)
5	(\$12; \$12)	(\$10, \$20)
6	(\$12; \$12)	(\$20, \$10)
7	(\$16; \$16)	(\$10, \$20)
8	(\$16; \$16)	(\$20, \$10)
9	(\$12; \$16)	(\$16, \$12)

Prior to playing each game, we asked subjects a series of 15 questions that tested their understanding of the instructions. Six understanding questions focused specifically on how votes translated to outcomes. Four understanding questions focused specifically on the fact that players were symmetrically uninformed in the SI Game. Five understanding questions focused specifically on the nature of the asymmetric information in the AI game.[6] All the instructions that subjects received are in Appendix G. A series of screen shots showing the understanding questions subjects faced are in Appendix H.

3.2 Experimental Procedures

The experiment is comprised of 5 parts. Part 1 is a simple decision-making task in which we introduce the notion of uncertainty, and which we use to inform us of subjects' risk attitudes. Part 2 introduces subjects to the voting structure that exists in the Main Game (i.e. the first option is implemented so long as it receives a single vote, while the second option is implemented if and only if both vote for it) but without uncertainty regarding the

[5]Round 9 is a "sanity check" to evaluate whether subjects paid attention to the values on their screens and whether subjects voted for the payoff-maximizing option when inequality and efficiency were the same in both options.

[6]We avoided introducing any elements that might lead subjects to "discover" that the informed player's vote carried information as to the payoffs in the risky option.

second option. Subjects played the SI game in Part 3, the AI game in Part 4, and ended with the Dictator games in Part 5. The order of rounds within each game was randomly determined at the subject level.

In each session, subjects received printed instructions for each part after they had completed the previous part, and those instructions were read aloud each time. Subjects could advance rounds within each part at their own pace, but the experiment advanced from part to part at the pace of the slowest subject. Subjects received no feedback as to their own or anyone else's choices. We conducted four sessions for a total of 86 subjects. Each session lasted about 50 minutes. This experiment took place in the Laboratory for Experimental Management and Auctions (LEMA) at Penn State University in the Spring of 2019.

In terms of payments, at the very start of each session, subjects were told that in addition to their $7 show-up fee, they would be paid for one part of the experiment only. We divided the understanding questions described above into three groups and attached them to Part 2 (where we introduce the voting structure), Part 3 (where subjects play the SI game) and Part 4 (where subjects play the AI game). Subjects were also told that if Part 2 or Part 3 or Part 4 was randomly chosen to count for payment, then they would be paid either for one randomly selected round in that part or for the understanding questions of that part. If the understanding questions were randomly chosen to count for payment, then they would earn $10 if they answered *all* questions of that part correctly; otherwise, they earned only $2. Average earnings were $15.

Because Parts 1 and 2 were primarily included to help subjects understand the AI game, we provide more details on those parts and the choices that subjects made in those parts in Appendix A. Therefore, the following section will focus on the AI game, as well as on behavior in the SI and Dictator games.

4 Results

We first describe behavior in the Asymmetric Information (AI) game and investigate whether, for a given value of the safe option, subjects in the role of the uninformed voter are more inclined to vote for the risky option when payoffs are positively correlated than when payoffs are negatively correlated. We then compare behavior across games in the HH treatment to distinguish the asymmetric-information rationale for this behavior from other confounds. Unless otherwise stated, all our claims are the results of within-subject analyses and the p-values we report correspond to Wilcoxon matched-pairs signed-rank tests.

Table 3: Aggregate results: fraction choosing the risky option in the HH treatment.

Round	Safe Option	Risky Option	AI Game (uninformed)	SI Game	Dictator Game
12N	($12; $12)	($10, $20) or ($20, $10)	47.7%	77.9%	72.1%
12P	($12; $12)	($10, $10) or ($20, $20)	86.0%	88.4%	82.6%
16N	($16; $16)	($10, $20) or ($20, $10)	1.2%	3.5%	0%
16P	($16; $16)	($10, $10) or ($20, $20)	32.6%	8.1%	7.0%

Table 3 displays aggregate data of subjects' choices. The fourth column shows the fraction of times subjects voted for the risky option when in the role of the uninformed voter in the AI game. The fifth column shows the same statistic but in the SI game, where both subjects are uninformed. The sixth column looks at the same behavior in a Dictator game, where a single uninformed subject chooses between the safe and risky options, knowing that her choice determines outcomes for both her and her partner.[7]

4.1 Do Subjects Account for Selection Effects?

At the aggregate level, subjects appear to respond to asymmetric information as predicted by the theoretical framework described in Section 2. In particular, we compare the number reported in the fourth column of Table 3 across the 12N and 12P Rounds, and then across the 16N and 16P Rounds. Within each of these pairs of rounds, the value of the safe option is held fixed and the only change is whether the risky option is negatively or positively correlated.

When the safe option is $12 and the outcomes from the risky option are negatively correlated, subjects choose the risky option 47.7% of the time compared with 86% of the time when they are positively correlated. When the safe option is $16, these numbers are 1.2% and 32.6%, respectively. For a given value of the safe option, the differences in behavior across positively and negatively correlated risky options are both large in magnitude and are statistically significant: whether the safe option is $12 or $16, subjects are significantly more likely to choose the risky option over the safe option when payoffs

[7]Appendix C shows subjects' choices in these as well as in the five other Dictator Games as shown in Table 2.

from the risky option are positively correlated than when payoffs are negatively correlated ($p < 0.001$ in both sets of comparisons).

To assess the degree to which subjects are reacting to asymmetric information in the AI game, we compare behavior in the AI with that of the SI games, where both players are symmetrically uninformed. Since the only distinction between these two games is in whether information is asymmetric, a change in subjects' behavior across these games is strong evidence that subjects are reacting to its presence. In particular, comparing the behavior of uninformed players in the AI and SI games, we should observe at least one of the following behaviors for a particular value of the safe option: (1) when the risky option is negatively correlated, a decrease in the fraction that vote for the risky option from the SI game to the AI game; (2) when the risky option is positively correlated, an increase in the fraction that vote for the risky option from the SI game to the AI game. Whether both or only one of these occurs depends on how risk aversion impacts choices in the SI game. Regardless of risk aversion however, the difference-in-differences across correlation structures for a given value of the safe option should be larger in the AI game than in the SI game.

We find substantial differences in behavior across the AI and SI games in line with these predictions, both at the round level, and when we compare difference-in-differences across correlation structures. For example, in the 12N round we see that subjects are far less likely to choose the risky option when information is asymmetric than when it is symmetric (47.7% versus 77.9% — $p < 0.001$). In parallel, in the 16P round, subject are far more likely to vote for the risky option when information is asymmetric than symmetric (32.6% versus 8.1% — $p < 0.001$). Both of these patterns are in line with the comparative predictions. Also demonstrating the impact of asymmetric information is the differences-in-differences in behavior across the 12N and 12P rounds as well as across the 16N and 16P rounds when we compare both games. Both those differences are much larger in the AI game than in the SI game: 38.3% versus 10.5% when $s = \$12$ ($p < 0.001$) and 31.4% versus 4.6% $s = \$16$ ($p < 0.001$).

While the theory matches qualitative predictions both within the AI game as well as across the AI and SI games, we do see significant departures from the point predictions in Section 2. If we look across all of the choices, 20.9% of subjects in the AI game behave according to all of the theoretical predictions, voting for the safe option in *both* negatively correlated rounds *and* voting for the risky option in *both* positively correlated rounds.[8]

[8] We measure consistency in all rounds because subjects answer only 4 questions in the asymmetric information game. We depict the full distribution of the number of deviations from Proposition 1 in Appendix D.

Among the subjects who do not fully conform to our predictions of Proposition 1, we identify differences in how consistently they conform in the positively and negatively correlated rounds.[9] The fraction of subjects who vote for the safe option in both of the negatively correlated rounds (the 12N and 16N rounds) is 52.3%, while the fraction of subjects who vote for the risky option in both of the positively correlated rounds (the 12P and 16P rounds) is lower at 30.2% ($p = 0.001$). These findings show that a greater fraction of subjects account for adverse selection rather than advantageous selection.

What else might be guiding subjects' choices? A poor understanding of our instructions does not appear to be a reason for the departures from theoretical predictions that we observe by some subjects.[10] In Sections 4.2 and 4.3 we discuss the degree to which the behavior that we observe can be explained by social preferences, strategic uncertainty, and failures of contingent reasoning.

4.2 The Role of Social Preferences

In this section, we explore the degree to which social preferences can explain behavior. Two leading theories of social preferences that could appear to play a role in our study are *aversion to ex post inequality* (e.g. Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) and *preferences for efficiency* (e.g. Charness and Rabin, 2002; Engelmann and Strobel, 2004). In both cases, our evidence suggests these theories do not fully explain the behavior that we observe in our experiments.

Aversion to ex post inequality: If subjects dislike ex post inequality, then this offers a rationale for them to choose the safe option when the risky option is negatively correlated but not when the risky option is positively correlated. Therefore, it offers a theoretically relevant confound because it has predictions that are identical to those of adverse and advantageous selection in the AI game.

[9] We observe no statistically significant differences in the proportion of subjects who conform to our theoretical predictions in the rounds in which $s = \$12$ and those in which $s = \$16$. Indeed, 39.5% of subjects behave according to our theoretical predictions in both rounds where $s = \$12$, and 32.6% do so when $s = \$16$ ($p = 0.239$).

[10] Recall that subjects faced a series of 15 questions that tested their understanding of the instructions. These questions were spread over the various Parts of the instructions. The median number of incorrect answers in the understanding questions is zero and the mean is 0.84 out of 15 questions. Both Chi Squared and Fisher exact tests show that the distribution of incorrect answers in the understanding questions among subjects who do not conform to Proposition 1 is no different than among those who do ($p = 0.808$ and $p = 0.959$, respectively). Further, subjects who answer all understanding questions perfectly are no more likely to conform to the predictions of Proposition 1 compared to those subjects who make at least one mistake in those understanding questions ($p = 0.411$). Thus we cannot attribute deviations from our theoretical predictions to confusion.

We find aversion to ex-post inequality may apply to only a few subjects. To see why, let us turn to the SI and Dictator games where neither player knows the payoffs of the risky option beyond its correlation structure. In both the SI and Dictator games, a large majority of subjects' decisions do not depend on whether the risky option's outcomes are negatively or positively correlated, even controlling for the amount of the safe option. Indeed, at the aggregate level, we see in Columns 5 and 6 of Table 3 that the differences between the 12N and 12P rounds, and between the 16N and 16P rounds, are not large in magnitude. The fractions of overlap between the 12N and 12P, and between the 16N and 16P rounds in the SI game are 84.9% and 93.0%, respectively. The corresponding fractions are 80.2% and 93.0% in the Dictator games.[11] At the individual level, if some subjects' choices are largely guided by aversion to ex-post inequality, then these subjects should behave according to our theoretical predictions in the AI game (though not necessarily due to selection) *and* play identically in the SI game. None of our subjects make choices that follow this pattern. Thus we rule out aversion to ex post inequality as a driver of behavior.

Preferences for efficiency: Similarly, we find that preferences for efficiency may apply to only a limited number of subjects. If subjects are largely motivated by the size of the total surplus, then we should see behavior that differs significantly from the theoretical predictions of Section 2. For example, when the safe option is \$12, then a subject with preferences for efficiency may, depending on how much she values efficiency relative to her own payoff, choose the risky option when it is negatively correlated, even when she is informed that the risky option lowers her own payoff. If the safe option is \$16, then such a subject may never choose the the risky option when it is negatively correlated, even if she is informed that the risky option increases her own payoff. We find that none of our subjects behave in a way that is consistent with preferences for efficiency across all rounds in the AI and Dictator game.[12] Even if we focus on the $s = \$12$ rounds separately, we find that at most 5 of our subjects behave in a way that is consistent with preferences for

[11] Subjects who do make different decisions across those rounds are generally more likely to favor the risky option when outcomes are positively correlated than when they are not (the p-values comparing the 12N and 12P rounds, as well as the 16N and 16P rounds in the SI and Dictator games are 0.013, 0.103, 0.029, 0.083).

[12] 14 subjects make decisions consistent with preferences for efficiency when informed (note that subjects do not see all the scenarios when informed, and some subjects only saw "advantageous" risky choices) and 8 subjects make decisions consistent with preferences for efficiency when uninformed. The intersection of those two groups represents 3 subjects. In addition, using behavior in the relevant rounds of the Dictator game, we find that none of those 3 subjects make the same efficient choices (these are Rounds 5, 6, 7 and 8 in Table 2).

efficiency, and in the $s = \$16$ rounds, only 6 of our subjects do so. Thus, it appears that the degree to which subjects in our experiment are motivated by efficiency is minimal.[13]

4.3 Strategic Uncertainty and Failures of Contingent Reasoning

Proposition 1 shows that if players are selfish, the unique strategy profile surviving two rounds of elimination of weakly dominated strategies involves the uninformed player choosing the safe option when payoffs are negatively correlated and the risky option when payoffs are positively correlated. While this logic may appear straightforward from the perspective of game theory, it involves two cognitive demands. First, it requires subjects to be confident that players behaving as informed voters do not choose weakly dominated actions. An uninformed Bob must attribute sufficiently high probability to the informed Alice choosing what is best for her that it rationalizes the equilibrium choice. This is an issue of strategic uncertainty. Second, it requires subjects to attend to a potentially non-salient feature of the game, namely that one's vote matters only when the other player is voting for the risky option. This is an issue of contingent reasoning. We investigate both of these below.

Let us first look at whether subjects are best-responding to the empirical distribution of play in the experiment. If it appears that a large fraction of subjects are not doing so, then this behavior suggests that subjects' behavior may be rationalized by strategic uncertainty, i.e., incorrect beliefs about the behavior of others. The first two columns in Table 4 show the possible rounds that the informed players saw, with the informed players' payoffs listed first.[14] The third column shows the fraction of informed players who choose the option with the payoff in bold. The fourth column shows the (ex ante) expected payoff for the uninformed player of choosing the risky option, given the empirical distribution of the informed players' choices.

We see that subjects who know the realized payoff of the risky option do not necessarily vote for the option that maximize their payoffs. When the safe option is \$12 and the risky option has negatively-correlated outcomes, 19.2% of informed subjects choose the safe option when they would have benefited from the risky option, and 9.4% choose the risky option despite it lowering their payoffs relative to the safe option. We see analogous behavior when the safe option is \$16 and the risky option has negatively correlated payoffs, but see relatively fewer departures from theoretical predictions when the risky option is positively correlated.

[13] We do not claim that such preferences do not exist. Rather that the marginal rates of substitution between one's own payoff and the social surplus may be such that, with our parameters, we don't observe such preferences, and thus they do not explain our subjects' behavior.

[14] Players did not see all of these rounds but only one in each pair of rows depending on the coin flip.

Table 4: Rationalizing "Mistakes": Expected Payoffs Given Empirical Distribution.

Round	Safe Option	"Risky" Option[a]	Fraction of Informed Players Choosing the "Risky" Option	Expected Payoff of Voting for the Risky Option Given Empirical Distribution[b]
12N	(\$12; \$12) (\$12; \$12)	**(\$10, \$20)** ~~or (\$20, \$10)~~ ~~(\$10, \$20) or~~ **(\$20, \$10)**	9.4% 81.8%	\$11.6
12P	(\$12; \$12) (\$12; \$12)	**(\$10, \$10)** ~~or (\$20, \$20)~~ ~~(\$10, \$10) or~~ **(\$20, \$20)**	2.4% 97.7%	\$15.9
16N	(\$16; \$16) (\$16; \$16)	**(\$10, \$20)** ~~or (\$20, \$10)~~ ~~(\$10, \$20) or~~ **(\$20, \$10)**	0% 82.2%	\$13.5
16P	(\$16; \$16) (\$16; \$16)	**(\$10, \$10)** ~~or (\$20, \$20)~~ ~~(\$10, \$10) or~~ **(\$20, \$20)**	2.4% 97.7%	\$17.9

[a]The ex-ante probability of either particular outcome was 50% but the informed player knew the outcome.

[b]This is for the uninformed voter given the choices/mistakes the informed voter makes empirically.

Inspecting the expected payoff for a subject in the role of the uninformed voter who votes for the risky option given the empirical distribution of play, we see that if such a subject had correct beliefs about the behavior of informed subjects, her decisions should coincide with the predictions from Section 2. Since we noted that only 20.9% of subjects followed these equilibrium predictions exactly, we do see evidence suggestive of strategic uncertainty, which motivates designing a treatment that eliminates strategic uncertainty, which we describe in Section 5. One interesting pattern that we note here is that there are relatively fewer departures from our theoretical predictions when these departures come at a higher cost.

Turning to the other cognitive demand, we investigate the degree to which subjects fail to apply contingent reasoning. Subjects who fail to apply contingent reasoning, should make the same choices in the AI and SI games since they are not thinking about the inference they should draw from being pivotal. Among those subjects who don't play the equilibrium strategies of Section 2, we see that slightly over half (57.4%) behave identically across the AI and SI games.[15]

[15]We note that our understanding questions in this treatment were deliberately designed to focus on the mechanics of the game and to avoid hinting that subjects should think about contingencies. As such, we cannot use the answers to these questions to assess the degree to which subjects fail or succeed in applying contingent reasoning.

5 The Human-Robot Treatment: Design and Results

To assess the importance of strategic uncertainty and failures of contingent reasoning, we conduct a "Human-Robot" (HR) treatment. Instead of being paired with another human subject, each subject is paired with a robot player whose strategy is revealed ahead of time. By pairing subjects with a computerized non-human subject in the SI and AI games, and telling our subjects how it had been programmed, we effectively remove issues of strategic uncertainty that potentially affected behavior in the main treatment.[16] An additional 82 subjects participated in the HR treatment. Below we detail how the HR treatment differs from our earlier HH treatment.

Symmetric Information Game: The parameters in the Symmetric Information game of the HR treatment were identical to those in the HH treatment. The instructions closely followed those in the HH treatment, except that subjects were now matched with a robot player that earned "virtual (imaginary) dollars" that "had no impact on you or anyone else at any point, ever." In the SI game, the robot player was programmed to always vote for the risky option. To closely match the understanding questions across treatments, subjects were told how the robot player was programmed only after they answered the understanding questions related to the mechanisms of the SI game. Directly following this information, subjects were asked to confirm they understood how the robot was programmed via one additional understanding question.

Asymmetric Information Game: In the Asymmetric Information game in the HR treatment, the robot player was always in the role of the informed voter and our subjects only participated in the role of the uninformed voter. The robot player was programmed to always vote for the option that gave it the highest amount of virtual (imaginary) dollars, and this was made known to the human subjects. The instructions in this treatment closely paralleled those in the HH treatment, as did the understanding questions.

Contingent Reasoning Questions and Asymmetric Information (2) Game: To evaluate subjects' ability to do contingent reasoning, we designed a new part following the AI game.[17] Subjects first answered a series of "contingent reasoning" (CR) questions, all of which pertained to the AI game they had just played (examples of these questions are in Section 5.3). These CR questions did not explain contingent reasoning to the subjects,

[16]It also removes social preferences, but as we concluded in our analysis of the HH treatment, these appear to play only a limited role in our experiment.

[17]This took the place of the Dictator game of the HH treatment.

but instead were designed to "nudge" subjects towards paying attention to contingencies. Following the CR questions, subjects again played against the robot players in a repetition of the AI game, which we call the AI(2) game. The CR questions permit us to match behavior in the AI game with subjects' abilities to answer questions on contingent reasoning, and then see whether such questions have a nudging effect in the AI(2) game.

We begin our analysis by comparing behavior in the AI and SI games in the HR treatment. We then compare behavior in these two games across the HH and HR treatments, and assess the degree to which strategic uncertainty influences behavior. Finally, we explore subjects' potential to reason about contingencies by evaluating their responses to the CR questions as well as behavior in the AI(2) game. Unless otherwise noted, the p-values associated with between-subjects comparisons across treatments are the result of two-sided tests of proportions, and the p-values associated with within-subject comparisons in the HR treatment are the result of Wilcoxon matched-pairs signed-ranks tests.

5.1 Aggregate Results in the HR Treatment

We present the aggregate data of the HR treatment in Table 5.

Table 5: Aggregate results: fraction choosing the risky option in the HR treatment.

Round	Safe Option	Risky Option	Asymmetric Information	Symmetric Information	Asymmetric Information(2)[a]
12N	(\$12; \$12)	(\$10, \$20) or (\$20, \$10)	23.2%	84.2%	22.0%
12P	(\$12; \$12)	(\$10, \$10) or (\$20, \$20)	81.7%	87.8%	90.2%
16N	(\$16; \$16)	(\$10, \$20) or (\$20, \$10)	2.4%	2.4%	1.2%
16P	(\$16; \$16)	(\$10, \$10) or (\$20, \$20)	46.3%	4.9%	62.2%

[a] Restricting attention to subjects who answered all questions correctly would generate fractions of 18.6%, 91.5%, 0% and 71.2% respectively.

We observe sharp differences in behavior when comparing behavior within the AI game across the 12N and 12P rounds, as well as across the 16N and 16P rounds, consistent with subjects responding to selection effects ($p < 0.001$ in both cases). Overall, 40.2% of the subjects behave in a way that is consistent across all rounds with our theoretical predictions from Section 2. We also note a large difference in behavior when comparing the difference-in-difference between the 12N and 12P rounds, as well as between the 16N and 16P rounds, across the SI and AI games: 58.5% versus 3.6% when $s = \$12$ ($p < 0.001$) and 43.9% versus 2.5% when $s = \$16$ ($p < 0.001$).

Finally, we also see that the difference in behavior in terms of how much subjects respond to adverse and advantageous selection persists in the AI game. In fact, in the HR treatment, almost three quarters of our subjects (74.4%) vote for the safe option in both the 12N and 16N rounds, corresponding exactly to our theoretical predictions from Section 2. In other words, all but a quarter of the subjects account perfectly for adverse selection. The corresponding fraction who vote for the risky option in both the 12P and 16P rounds, where payoffs are positively correlated, is 42.7%. Thus, we see evidence both that a substantial fraction of our subjects account perfectly for selection and yet, a gap between adverse and advantageous selection remains among those who do not ($p < 0.001$).

5.2 The Impact of Strategic Uncertainty

To assess the role of strategic uncertainty we compare our results for the HR treatment with the HH treatment. Figure 1 shows the proportion of subjects choosing the risky option in the AI game for both the HH and HR treatment.[18] We illustrate the choices in the AI games for both the HH and HR treatment. Since we observe no statistical difference across the HH and HR treatments in the SI games, for simplicity we provide the average choices in the SI games across HH and HR treatments.[19][20]

We find that subjects respond more to selection effects when strategic uncertainty is removed. This is the case in the 12N and 16P rounds, which were the rounds in which a substantial fraction of subjects deviated from the theoretical predictions in the HH treatment. In the 12N round, the fraction of subjects who chose the risky option in the HH treatment was 47.7%, while it is 23.2% in the HR treatment where strategic uncertainty is eliminated ($p = 0.001$). In parallel, the fraction of subjects who chose the risky option in the 16P round of the HR treatment is 46.3%, up from 32.6% in the HH treatment ($p = 0.068$). Overall, in the HR treatment, 40.2% of subjects behaved according to the theoretical predictions in Section 2 for all rounds of the AI game. This fraction is significantly higher than in the HH treatment, where it was 20.9% ($p = 0.007$).

How much of the difference in behavior across treatments can we attribute to strategic

[18]Recall that according to Proposition 1, in the AI game, subjects should vote for the safe option in the negatively correlated rounds, and vote for the risky option in the positively correlated rounds.

[19]Indeed, the smallest p-value is 0.303 when comparing behavior in the four SI rounds across the HR and HH treatments. Thus we cannot reject the null that the answers to these questions come from the same population.

[20]The insignificant difference in subjects' behavior in the SI games in the HH and HR treatments provides further evidence that social preferences play only a limited role in our experiment.

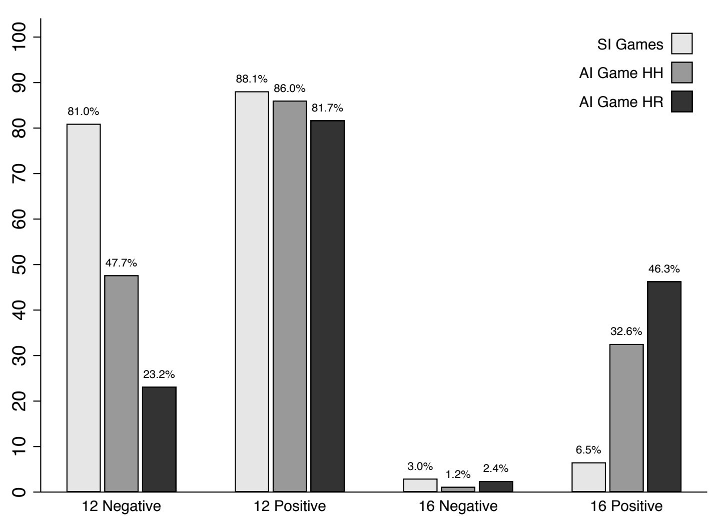


Figure 1: Fraction choosing the risky option in Human-Human and Human-Robot Treatments.

uncertainty, as opposed to difference in subject characteristics across treatments? We note that there are no discernible differences between the subjects in the HH and HR treatments in terms of their demographics or how well they understood the instructions.[21] Thus, neither of these explain the increase in the proportion of subjects whose choices are in line with theoretical predictions in the HR treatment. In addition, since removing strategic uncertainty has no impact on how subjects play the SI game (see previous sub-section), it also does not seem that subjects across treatments differed in their social preferences, or beliefs about their pivotality. So, we cautiously attribute treatment differences to strategic uncertainty and estimate that it accounts for about 25% of the deviations from the theoretical predictions that we observed in the HH treatment.

5.3 Contingent Reasoning

A plausible conjecture is that subjects who continue to depart from our theoretical predictions even after strategic uncertainty is eliminated are those who simply cannot reason about contingencies. We test this conjecture in Part 5 of the HR treatment. After the AI game, subjects answer a series of questions that draw attention to the information conveyed in the robot player's vote, and thus contingent reasoning. These contingent reasoning questions (CR) take place before subjects play the AI game a second time. We

[21]For demographic data in our two treatments, see Section E. Restricting attention to the understanding questions in the two treatments that share a common structure, both Chi2 and Fisher exact tests fail to reject that the distribution of mistakes are from the same population (the p-values are 0.797 and 0.940).

investigate how responses to these questions correlate with subjects' behavior in the AI game on the first iteration and how answering these questions influences behavior in the AI game on the subsequent play.

The first two CR questions assess whether subjects understand that the vote of the Robot player carried information on the coin flip. The remaining two assess whether they understood that this could impact their own payoff. An example of the former and latter are below, where the items in the square brackets correspond to the multiple choice answers the subjects faced.[22]

> *Given how the computer player was programmed in Part 4, if the computer player votes for the option requiring 2 votes (the option on the right), what does that tell you about the outcome of the coin flip? [That it landed on HEADS; that it landed on TAILS, it doesn't tell you anything about the outcome of the coin flip]*

> *Given how the computer player was programmed in Part 4, if the computer player votes for the option requiring 2 votes and you vote for that option too, how much will you earn? [$15, $17, $20, You will earn $15 or $20 with equal chance of each.]*

We find that 89% of the subjects answer *all* of the CR questions correctly. Correlating these responses with behavior in the preceding asymmetric information game, we find that all subjects who behaved exactly according to our theoretical predictions in the preceding AI game answered each CR question correctly. Of the 11% of players who answered at least one CR question incorrectly, none played according to our theoretical predictions in the preceding AI game.

How does answering these CR questions affect subsequent play? The last column in Table 5 displays the fraction of subjects who vote for the risky option for each round of the AI(2) game. The fraction of subjects who behave according to the theoretical predictions increases in both positively correlated rounds ($p = 0.035$ when the safe option is $12, and $p = 0.003$ when the safe option is $16), and is statistically no different in the two negatively correlated rounds ($p > 0.100$ in both cases). Overall, 57.3% of subjects in the AI(2) game make all their choices in a way that is consistent with the theoretical predictions; this fraction is statistically higher than that in the AI game played before the CR questions (40.2%) ($p < 0.001$). Thus, the CR questions help some subjects understand selection effects but a significant fraction of subjects continue to deviate from theoretical predictions.

[22] For reference, in the questions below, "Part 4" refers to the AI game. To explain the nature of the uncertainty, throughout the instructions we used the example of a fair coin flip that determined what the payoffs in the risky option would be if it was to be implemented.

Interestingly, a large fraction of these subjects appear to understand contingent reasoning when it is broken down into steps: of the 42.7% of subjects whose choices depart from our theoretical predictions in the AI(2) game, 77.1% actually answer all contingent reasoning questions correctly. These subjects show that they understand that the robot player's votes carries information about their own payoff, and yet make choices in the AI(2) game that lead to lower payoffs. Thus, these subjects show that they are able to correctly execute each step of contingent reasoning separately but do not piece together these steps when they subsequently play the AI(2) game.

The data show that the asymmetry in the degree to which subjects account for negative versus positive correlation persists in the AI(2) game. Looking at only decisions in negatively correlated rounds, 76.8% of subjects match our theoretical predictions while the analog for positively correlated rounds is 61.0%, which is significantly less ($p = 0.003$).

5.4 Role of Limited Strategic Thinking

So we see that even after removing strategic uncertainty, a gap remains in the degree to which subjects account for adverse and advantageous selection. This section explores whether models of limited strategic thinking, in the form of level-k or cursed equilibrium, can explain this gap. Our analysis suggests that it can theoretically do so, but would require a high degree of risk-aversion that is not supported in our data.

For a generic subject, denote her utility from wealth w by $u(w)$. We assume that u is strictly increasing and continuous.[23] We allow for both risk-averse and risk-seeking behavior, but assume that risk attitudes are stable across the wealth levels that we study. Accounting for adverse selection but not advantageous selection corresponds to a player choosing the safe option in both negatively and positively correlated rounds when he is uninformed (i.e., in the situation of Bob in Section 2). Without making any further assumptions on the utility function, we study when this is possible for both level-k and cursed equilibrium behavior.

Level-k: Let us begin with level-k. Consider a random-$L0$ specification in which $L0$ votes for safe with probability p and for risky with probability $(1-p)$ where p is in $(0, 1)$, regardless of whether the $L0$ player is informed or uninformed.[24] By Proposition 1, if the player is $L2$ or above, then the player must choose the risky option whenever payoffs

[23]Since we apply these models to our HR treatment, we assume that subjects exhibit no social preferences to their robot partners.

[24]Most formulations (Crawford, Costa-Gomes, and Iriberri, 2013) assume that a random-$L0$ player uniformly randomizes, but this is unnecessary for the conclusion that we draw here.

are positively correlated. Thus, an uninformed player chooses the safe option in both negatively and positively correlated rounds only if she is either $L0$ or $L1$. An $L1$-player chooses the safe option when the value of the safe option is \$12 and he is uninformed if and only if

$$u(12) \geq pu(12) + (1-p)\left(\frac{1}{2}u(10) + \frac{1}{2}u(20)\right).$$

The LHS is the payoff from choosing S and the RHS is the expected payoff from choosing R, assuming that one's opponent is a random-$L0$ player. Rearranging the above inequality indicates that such a player prefers obtaining \$12 for sure to a symmetric lottery between \$10 and \$20. In our risk-elicitation task, we see that only 2.4% of subjects exhibit such preferences (Table 9 in Appendix B). Thus, to explain why subjects may account for adverse but not advantageous selection, level-k requires a degree of risk aversion beyond that which we see in our data.

Cursed Equilibrium: We turn to cursed equilibria (Eyster and Rabin, 2005) and let χ in $[0,1]$ denote the degree of cursedness of a player. In a cursed equilibrium, a player in the role of the uninformed voter has the correct marginal beliefs about the behavior of her partner but does not sufficiently appreciate how that behavior is affected by private information. A player chooses a \$12 safe option when payoffs are negatively correlated if

$$u(12) \geq (1-\chi)\left(\frac{1}{2}u(12) + \frac{1}{2}u(10)\right) + \chi\left(\frac{1}{2}u(12) + \frac{1}{4}u(10) + \frac{1}{4}u(20)\right), \tag{1}$$

where the LHS is the payoff from choosing S and the RHS is the expected payoff from choosing R for a χ-cursed player. Similarly, a player chooses a \$16 safe option when payoffs are positively correlated if

$$u(16) \geq (1-\chi)\left(\frac{1}{2}u(16) + \frac{1}{2}u(20)\right) + \chi\left(\frac{1}{2}u(16) + \frac{1}{4}u(10) + \frac{1}{4}u(20)\right). \tag{2}$$

We show that if the certainty equivalent for a 50-50 lottery on \$10 and \$20 weakly exceeds \$14, there is no value of χ that satisfies both (1) and (2). (The proof is in Appendix F.)

Proposition 2. *If $u(14) \geq (u(10) + u(20))/2$, then there is no value of χ for which (1) and (2) are simultaneously satisfied.*

Almost 90% of our subjects have a switching point of \$14 or above (Table 9 in Appendix B), and over 70% of our subjects have a switching point of \$15 or above. Moreover, in both of these sub-samples, subjects continue to account more for adverse selection than advantageous selection ($p < 0.001$ and $p = 0.002$, respectively).

6 The Feedback Treatments: Design and Results

After our various treatments, we are therefore left with a puzzle: why are subjects more likely to account for adverse selection than advantageous selection? We hypothesize that a contributing factor to this gap is that in everyday life, people obtain payoff feedback from the choices they make but rarely observe counterfactuals. This limitation in feedback has a differential effect on behavior across settings with strategic selection. Let us explain why.

Uninformed individuals who repeatedly choose a risky option when outcomes are negatively correlated would see that they are better off from choosing the safe option. This feedback allows them to learn from their mistakes so that these mistakes do not persist in the long-run. On the other hand, if uninformed individuals repeatedly choose the safe option when payoffs from the risky options are positively correlated, they do not observe what would have happened had they chosen the risky option instead. Hence, they do not learn from their mistake, and thus, such mistakes persist in the long-run.

We formalize this logic in the language of self-confirming equilibria.

Proposition 3.

(a) *If payoffs are negatively correlated, then in every weakly undominated self-confirming equilibrium, the uninformed player votes for the safe option.*

(b) *If payoffs are positively correlated, then there exists a weakly undominated self-confirming equilibrium in which the uninformed player votes for the safe option.*

Proposition 3 tells us that if payoffs are negatively correlated, incorrect beliefs about off-path behavior cannot rationalize departures from the predictions of weakly undominated Bayes-Nash equilibria (Proposition 1). In other words, learning dynamics would lead players to account for adverse selection. By contrast, if payoffs are positively correlated, then there exist beliefs about off-path behavior that can induce someone to choose differently from Bayes-Nash equilibria. Opportunities for feedback and learning do not mitigate this mistake because players do not obtain the payoff feedback that identifies the mistake. The proof for Proposition 3 is straightforward, and is in Appendix F.

One way to test whether this mechanism is at play is to see how subjects respond to feedback about counterfactuals and off-path histories. Proposition 3 has two implications.

First, it indicates that varying payoff feedback should have little effect if payoffs are negatively correlated but have a significant effect if payoffs are positively correlated. Second, the gap between positive and negatively correlated payoffs should reduce if subjects are given feedback about counterfactuals, but not if that information is absent.

We test these predictions in our two subsequent treatments, Partial Feedback (PF) and Full Feedback (FF). Each is identical to the HR treatment except for Part 4, where subjects play the AI game against a robot player multiple times but now obtain payoff feedback. The PF treatment resembles our description of everyday life: in the PF treatment, after each feedback round, a subject is reminded of how he voted and told what the payoffs would be if that round is selected for payment. Thus, a subject choosing the safe option does not learn the Robot's vote or the coin flip, and cannot deduce what would have happened had he chosen the risky option. By contrast, in the FF treatment, after each feedback round, a subject is reminded of his vote and told the result of the coin flip, how the computer voted, which of the two options was implemented for the round, and what his payoffs would be if that round is selected for payment. Thus, in the FF treatment, regardless of a subject's vote, that subject can deduce what the payoffs would have been had he voted differently.[25]

More specifically, in both the PF and FF treatments, Part 4 comprises 5 repetitions of each of the 12N, 12P, 16N and 16P rounds of the AI game described in Table 1 . Within each session, half of the subjects saw the 10 rounds of negatively correlated outcomes first. Within those 10 rounds, the fixed amount was either \$12 or \$16, each happening 5 times, in random order. These subjects then saw the 10 rounds of positively correlated outcomes, again where the fixed amount was \$12 or \$16 in random order. The other half of the subjects saw the positively correlated rounds first.[26] A total of 83 subjects participated in the FF treatment, and a total of 86 subjects participated in the PF treatment. After these feedback rounds, subjects face positively and negatively correlated payoffs, exactly as in Part 5 of the HR treatment. These Part 5 rounds involve no feedback.

Our analysis concerns how partial and full feedback in Part 4 affects behavior in Part 5, namely behavior in the subsequent rounds without feedback. Table 6 lists the fraction of subjects who choose the safe option in negatively correlated rounds and the risky option in positively correlated rounds. When payoffs are negatively correlated, the fraction choosing the safe option is not significantly different across the partial and full feedback

[25] These instructions, as well as screen shot examples showing what the feedback rounds looked like are in Appendix G.

[26] The transition from the negatively to positively correlated rounds (and vice versa) was seamless: subjects simply moved from one type of setting to the next without any announcement.

Table 6: *Fraction of subjects following theoretical predictions in Part 5.*

	Partial Feedback Treatment	Full Feedback Treatment
Both Negatively Correlated Rounds	77.9%	81.9%
Both Positively Correlated Rounds	62.8%	75.9%
All Rounds	55.8%	71.1%

treatments (77.9% and 81.9%, $p = 0.515$). However, when payoffs are positively correlated, significantly more subjects choose the risky option in the full feedback treatment compared to the partial feedback treatment (75.9% versus 62.8%, $p = 0.065$). Thus, we see evidence consistent with the implication of Proposition 3 that feedback about counterfactuals should be important if payoffs are positively correlated, but not if payoffs are negatively correlated.

Also in line with Proposition 3, we see that with partial feedback, subjects continue to react differently across the positively and negatively correlated rounds ($p = 0.009$). By contrast, with full feedback, the difference between the positively and negatively correlated rounds is no longer significant ($p = 0.166$). Thus, full feedback not only increases the fraction of subjects who behave according to the predictions of Proposition 1 in the positively correlated rounds but also closes the gap in behavior across positively and negatively correlated rounds, whereas with partial feedback, this gap remains.

Finally, we can also compare behavior here with that of our earlier no-feedback Human-Robot treatment to see how feedback influences behavior. We find that partial feedback does not significantly affect behavior: in the AI(2) game of the no-feedback HR treatments (described in Section 5.3), the proportion of subjects matching the predictions of Proposition 1 in negatively-correlated, positively-correlated, and all rounds are 76.8%, 61.0%, and 57.3% respectively. These proportions do not significantly differ from those of the partial feedback treatment (described in Table 6); the p-value exceeds 0.1 in each case. These proportions are, however, significantly different from those for the Full Feedback treatment for positively-correlated and all rounds ($p = 0.039$ and $p = 0.065$, respectively) but not for the negatively-correlated rounds ($p = 0.418$). This is consistent with the hypothesis that individuals may be learning how to cope with adverse selection from everyday experience, and thus, neither partial nor full feedback significantly influences their behavior in these settings. By contrast, in settings with advantageous selection, feedback has a significant effect on behavior when subjects observe the counterfactual.

Our analysis elucidates how the inability to observe counterfactuals biases learning so that people may learn to account for adverse selection but do not learn to account for advantageous selection. Our analysis does not shed light on the origin of these biases but helps us to understand how these biases persist, and why everyday learning may not eliminate them. Initial biases that may cause individuals to distrust others in settings with advantageous selection (such as "zero-sum thinking") may persist despite opportunities to learn from experience. By contrast, when individuals are biased in favor of trusting others in settings with adverse selection, everyday learning from the actions that one takes is sufficient to correct such biases.

7 Conclusion

We study how people respond to adverse and advantageous selection using a simple two-person game where asymmetrically informed subjects choose between a risky option and a safe option. We vary whether payoffs from the risky option are negatively correlated (inducing adverse selection) or positively correlated (inducing advantageous selection). To isolate the role of asymmetric information from other confounds, these subjects also play a game that is otherwise identical but where both players are symmetrically uninformed.

We find that uninformed subjects are more likely to choose the risky option when payoffs are positively correlated than when payoffs are negatively correlated. These differences do not arise when players are symmetrically uninformed, indicating that subjects respond to asymmetric information. But we also see departures from theoretical predictions. In particular, subjects are more likely to account for adverse rather than advantageous selection. Our subsequent treatments help us diagnose why we see these departures.

The second treatment studies how strategic uncertainty—uncertainty about the behavior of others—influences choice. We pair subjects with a computerized robot player whose strategy is known. A cross-treatment comparison shows that strategic uncertainty explains up to a quarter of the departures from theoretical predictions in the first treatment (when subjects were paired with each other). In the second treatment, we also ask a number of questions that explore subjects' understanding of contingent reasoning. Answering these questions affects behavior for a significant fraction of our subjects, indicating that contingent reasoning can be learned. However, we also find that a non-trivial fraction of subjects demonstrate an excellent understanding of contingent reasoning when asked questions about it but fail to implement that knowledge in a strategic setting even after those questions. These subjects appear to understand each element of contingent reasoning separately but do not piece them together on their own.

Our third and fourth treatments explore whether the gap in the degree to which subjects account for adverse versus advantageous selection relates to the inability to learn about counterfactuals. We vary payoff feedback in our experimental design and see how much of an effect feedback about counterfactuals can play. We find that it closes the gap: full feedback leads to no significant differences in how well subjects account for adverse versus advantageous selection whereas a significant gap remains with partial feedback.

To summarize, people do account for asymmetric information but the degree to which they do so is contextual. When there is reason for "distrust"—such as in settings of negatively-correlated payoffs—people do not let better informed partners make the final choice. But when there is reason to trust those who are better-informed—because payoffs are positively correlated—people fail to capitalize on these gains.

We view these results to be germane to political and social interactions. They suggest a potential mechanism for the prevalence and persistence of "zero-sum thinking" noted in social psychology: people learn to distrust others because mistakes from zero-sum games are self-correcting whereas those from settings with common interests are not. Such behavior may have direct consequences for political behavior and elections. To the extent that voters perceive there to be significant political polarization (Levendusky and Malhotra, 2015; Alesina, Miano, and Stantcheva, 2020), our results suggest that voters are likely to be suspicious of the information possessed by others and unlikely to capitalize on gains that could come from advantageous selection.

References

Akerlof, G. A. (1970): "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 488–500.

Alesina, A., A. Miano, and S. Stantcheva (2020): "The Polarization of Reality," *American Economic Review Papers and Proceedings*.

Ali, S. N., M. Mihm, and L. Siga (2018): "Adverse selection in distributive politics," Working Paper.

Barron, K., S. Huck, and P. Jehiel (2019): "Everyday econometricians: Selection neglect and overoptimism when learning from others," WZB Discussion Paper.

Battaglini, M., R. B. Morton, and T. R. Palfrey (2010): "The Swing Voter's Curse in the Laboratory," *Review of Economic Studies*, 77, 61–89.

Bazerman, M. H. and W. F. Samuelson (1983): "I Won the Auction But Don't Want the Prize," *Journal of Conflict Resolution*, 27, 618–634.

Bolton, G. E. and A. Ockenfels (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 90, 166–193.

Brocas, I., J. D. Carrillo, S. W. Wang, and C. F. Camerer (2014): "Imperfect choice or imperfect attention? Understanding strategic thinking in private information games," *Review of Economic Studies*, 81, 944–970.

Carrillo, J. D. and T. R. Palfrey (2009): "The Compromise Game: Two-Sided Adverse Selection in the Laboratory," *American Economic Journal: Microeconomics*, 1, 151–81.

——— (2011): "No Trade," *Games and Economic Behavior*, 71, 66–87.

Charness, G. and D. Levin (2009): "The Origin of the Winner's Curse: A Laboratory Study," *American Economic Journal: Microeconomics*, 1, 207–36.

Charness, G. and M. Rabin (2002): "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817–869.

Crawford, V. P., M. A. Costa-Gomes, and N. Iriberri (2013): "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications," *Journal of Economic Literature*, 51, 5–62.

De Meza, D. and D. C. Webb (2001): "Advantageous Selection in Insurance Markets," *RAND Journal of Economics*, 249–262.

Engelmann, D. and M. Strobel (2004): "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments," *American Economic Review*, 94, 857–869.

Enke, B. (2019): "What you see is all there is," Working Paper.

Esponda, I. and E. Vespa (2014): "Hypothetical Thinking and Information Extraction in the Laboratory," *American Economic Journal: Microeconomics*, 6, 180–202.

——— (2018): "Endogenous Sample Selection: A Laboratory Study," *Quantitative Economics*, 9, 183–216.

——— (2019): "Contingent Preferences and the Sure-Thing Principle: Revisiting Classic Anomalies in the Laboratory," Working Paper.

Eyster, E. and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623–1672.

Fang, H., M. P. Keane, and D. Silverman (2008): "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy*, 116, 303–350.

Feddersen, T. and W. Pesendorfer (1996): "The Swing Voter's Curse," *American Economic Review*, 86, 408–424.

Fehr, E. and K. M. Schmidt (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114, 817–868.

Fudenberg, D. and D. K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 523–545.

——— (1997): "Measuring Players' Losses in Experimental Games*," *The Quarterly Journal of Economics*, 112, 507–536.

Fudenberg, D. and E. Vespa (2019): "Learning Theory and Heterogeneous Play in a Signaling-Game Experiment," *American Economic Journal: Microeconomics*, 11, 186–215.

Guarnaschelli, S., R. D. McKelvey, and T. R. Palfrey (2000): "An Experimental Study of Jury Decision Rules," *American Political Science Review*, 94, 407–423.

Kagel, J. H. and D. Levin (1986): "The winner's curse and public information in common value auctions," *The American economic review*, 894–920.

Levendusky, M. S. and N. Malhotra (2015): "(Mis)perceptions of Partisan Polarization in the American Public," *Public Opinion Quarterly*, 80, 378–391.

Magnani, J. and R. Oprea (2017): "Why Do People Violate No-Trade Theorems? A Diagnostic Test," Working Paper.

Martínez-Marquina, A., M. Niederle, and E. Vespa (2019): "Failures in Contingent Reasoning: The Role of Uncertainty," *American Economic Review*, 109, 3437–74.

Meegan, D. V. (2010): "Zero-Sum Bias: Perceived Competition Despite Unlimited Resources," *Frontiers in Psychology*, 1, 191.

Milgrom, P. and N. Stokey (1982): "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17–27.

Rothschild, M. and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90, 629–649.

Różycka-Tran, J., P. Boski, and B. Wojciszke (2015): "Belief in a Zero-Sum Game as a Social Axiom: A 37-Nation Study," *Journal of Cross-Cultural Psychology*, 46, 525–548.

Sonsino, D., I. Erev, and S. Gilat (2002): "On rationality, learning and zero-sum betting—an experimental study of the no-betting conjecture," Working Paper.

Online Appendices

The appendix is organized as follows. Appendix A provides summary statistics for Parts 1 and 2 in the HH treatment, where Part 1 involved individual choices over lotteries and Part 2 involved a voting game without uncertainty about payoffs. Appendix B summarize the corresponding choices for Part 1 in the HR treatment. Appendix C provides a summary of all choices in the Dictator game, which was Part 5 of the HH treatment. Appendix D summarize the distribution of the number of deviations from our theoretical predictions in Proposition 1 for all of our treatments (HH, HR, PF and FF). Appendix E summarizes demographic information on our subject pool in each of our treatments, and Appendix F provides the proofs of Propositions 2 and 3. Full experimental instructions with screenshots are given in Appendix G.

A Analysis of Parts 1 and 2 in the HH treatment

A.1 Part 1 - lottery questions

In Part 1, subjects faces a series of 9 rounds of an individual decision-making task. In each round subjects had the choice between two Options. Option A was a fixed amount (that varied from round to round). Option B was a lottery that paid \$10 with 50% chance, and \$20 with 50% chance. For each subject, the fixed amounts were drawn randomly without replacement from the following list $\$11, \$12, ..., \$18, \$19$. After Part 1 was over, subjects were informed that from that point onwards, they would be matched into pairs with another player in the room.

Table 7 shows the fraction of subjects choosing the lottery for each round. A large majority of subjects (88.4%) have a single switching point. That is, there is a fixed amount above which they always forgo the lottery and below which they always choose the lottery. Almost all subjects prefer the lottery over the fixed amount of \$12, and almost all prefer \$16 to the lottery.

Table 7: Part 1 choices.

Round[a]	Option A	Option B[b]	Fraction choosing the lottery	Switching point (assuming single SP)
1	11	Lottery	97.7%	0%
2	12	Lottery	94.2%	5.3%
3	13	Lottery	90.7%	4%
4	14	Lottery	62.8%	26.3%
5	15	Lottery	15.1%	51.3%
6	16	Lottery	4.7%	11.8%
7	17	Lottery	1.2%	1.3%
8	18	Lottery	2.3%	0%
9	19	Lottery	0%	0%
Fraction (number) of subjects with a single switch point: 88.4% (76)				

[a] The order of rounds was randomly determined for each subject in each session, so was which option appeared on the left or right of the screen.

[b] In all rounds, the lottery paid \$10 with 50% chance and \$20 with 50% chance.

A.2 Part 2 - voting, no uncertainty

In Part 2, subjects play 6 rounds of a game in which each player in a pair has to vote for one of two options that determine outcomes for both players. Just as in the Main Game, Option A is implemented so long as it receives at least one vote, while Option B requires two votes to be implemented. Unlike in the Main Game, Option B here consists of a fixed and known allocation. Table 8 presents the parameters that subjects faced in each round as well as the fraction of subjects who chose Option B in each of those rounds. The order in which rounds were presented to the subjects was randomly determined and thus varied from subject to subject.

Table 8: Part 2 choices.

Round[a]	A (1 vote)	B (2 votes)	Fraction Choosing Option B
1	(\$12 ; \$12)	(\$10 ; \$20)	0.198
2	(\$12 ; \$12)	(\$20 ; \$10)	0.686
3	(\$16 ; \$16)	(\$10 ; \$20)	0.023
4	(\$16 ; \$16)	(\$20 ; \$10)	0.419
5	(\$12 ; \$16)	(\$16 ; \$12)	0.872
6	(\$16 ; \$12)	(\$12 ; \$16)	0.058

[a]The order of rounds was randomly determined for each subject in each session.

B Analysis of Part 1 in the HR treatment

Part 1 was identical in the HH and HR treatments. Table 9 shows the fraction of subjects choosing the lottery for each round. A large majority of subjects (93.9%) have a single switching point. That is, there is a fixed amount above which they always forgo the lottery and below which they always choose the lottery. Almost all subjects prefer the lottery over the fixed amount of \$12, and almost all prefer \$16 to the lottery. In addition, over 70% of subjects have a switching point of \$15, and close to 90% have a switching point of \$14 or more.

Table 9: Part 1 choices.

Round[a]	Option A	Option B[b]	Fraction choosing the lottery	Switching point (assuming single SP)
1	11	Lottery	100%	0%
2	12	Lottery	97.6%	1.3%
3	13	Lottery	85.6%	9.1%
4	14	Lottery	70.7%	19.5%
5	15	Lottery	11.0%	59.7%
6	16	Lottery	2.4%	9.1%
7	17	Lottery	1.2%	1.3%
8	18	Lottery	1.2%	0%
9	19	Lottery	1.2%	0%
Fraction (number) of subjects with a single switch point: 93.9% (77)				

[a]The order of rounds was randomly determined for each subject in each session, so was which option appeared on the left or right of the screen.

[b]In all rounds, the lottery paid \$10 with 50% chance and \$20 with 50% chance.

C Part 5 statistics in the HH treatment

Part 5 in the HH treatment was a Dictator game, where a subject's choice determined her payoff and the payoff of her partner. The payoffs were symmetric under Option A and asymmetric under option B, and Table 10 shows the fraction of subjects choosing option B in each round.

Table 10: Fraction Choosing Option B in the Dictator Game in the HH treatment.

Round	Option A	Option B	Fraction Choosing Option B
1	(\$12; \$12)	(\$10, \$20) or (\$20, \$10)	72.1%
2	(\$12; \$12)	(\$10, \$10) or (\$20, \$20)	82.6%
3	(\$16; \$16)	(\$10, \$20) or (\$20, \$10)	0%
4	(\$16; \$16)	(\$10, \$10) or (\$20, \$20)	7.0%
5	(\$12; \$12)	(\$10, \$20)	14.0%
6	(\$12; \$12)	(\$20, \$10)	94.2%
7	(\$16; \$16)	(\$10, \$20)	0%
8	(\$16; \$16)	(\$20, \$10)	70.9%
9	(\$12; \$16)	(\$16, \$12)	98.8%

D Distribution of the Number of Deviations from Proposition 1

Figure 2 shows the number of deviations from Proposition 1 that subjects in the role of the uninformed player make in Part 4 of the HH and HR treatments, as well as in Part 5 of the PF and FF treatments (after the feedback rounds). The data show that of those subjects who are not fully consistent with these theoretical predictions, subjects are more likely to have two deviations rather than one deviation. We also see treatment effects from removing strategic uncertainty (in the Human-Robot Treatment) and introducing full feedback (in the Feedback Treatment).

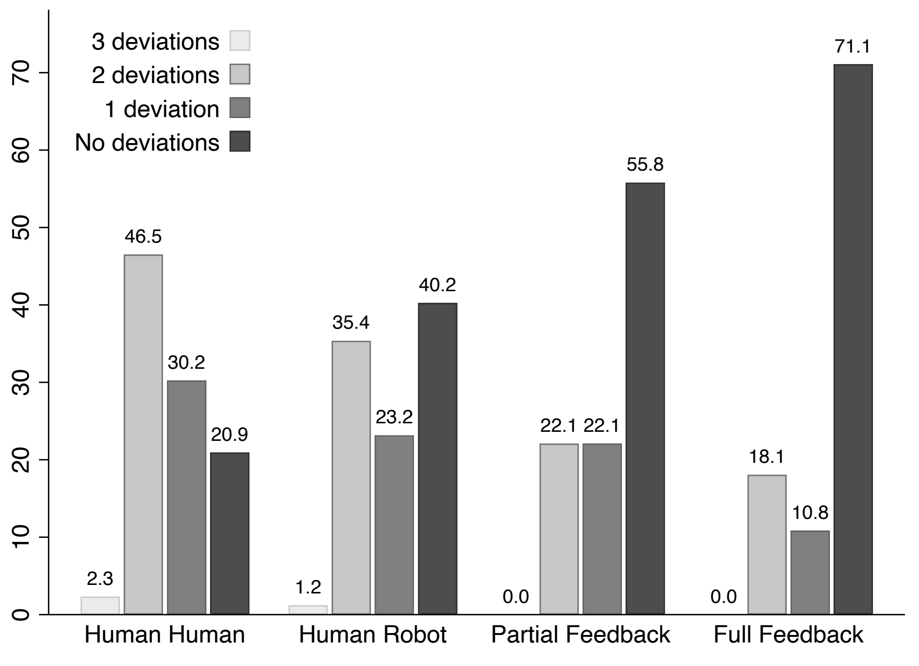


Figure 2: Distribution of the number of deviations from Proposition 1 by treatment.

E Demographic Information

Table 11 presents statistics on the demographic information that we collected from subjects via a questionnaire.

Table 11: Average demographic information.

	Human-Human Treatment	Human-Robot Treatment	Partial Feedback Treatment	Full Feedback Treatment
Female	55.8%	64.3%	64.0%	63.9%
Age	21.1	21.0	22.2	22.0
GPA	3.4	3.5	3.4	3.5
Nb Years at PSU	3.2	2.9	3.0	2.6
Nb. of subjects	86	82	86	83

A series of Fisher exact and Chi-squared tests (age, school/major at PSU), test of probability (female) and ranksum test (GPA) reject the hypotheses that subjects in the treatments come from different populations.[27] We find that demographic details (sex, age, GPA, or other observables) have no significant effect on behavior.

[27]In all treatments, the most represented major/school was Economics or Business followed by Science and Engineering. The fraction of subjects who are in those fields in each treatment represents 74.4%, 63.4%, 67.4% and 63.9%, respectively.

F Proofs of Propositions

Proof of Proposition 2 on p. 25. Without loss of generality, we can normalize utilities so that $u(10) = 0$ and $u(20) = 1$. With this normalization, (1) simplifies to $u(12) \geq \chi/2$ and (2) simplifies to $u(16) \geq 1 - \chi/2$. Let CE denote the certainty equivalent of a $50 - 50$ lottery on \$10 and \$20, which implies that $u(CE) = 1/2$. We now consider the three separate cases of risk-aversion, risk-neutrality, and risk-seeking behavior.

Let us first consider the case in which u is strictly concave and $CE \in [14, 15)$. Then it follows from the concavity of u that

$$u(CE) - u(12) > u(16) - u(CE),$$

which combined with $u(CE) = 1/2$ implies that $u(16) < 1 - u(12)$. But then if $u(12) \geq \chi/2$ because u and χ satisfy (1), it must be that $u(16) < 1 - \chi/2$, which implies that (2) cannot be satisfied.

Now consider the case of risk-neutral preferences. Given our normalization, it must be that $u(16) = 3/5$ and $u(12) = 1/5$, which implies that (1) and (2) cannot be simultaneously satisfied.

Finally, consider the case in which u is strictly convex and that CE is in $(15, 20]$. Because u is strictly convex,

$$\frac{u(16) - u(CE)}{16 - CE} < \frac{u(20) - u(CE)}{20 - CE} = \frac{1}{2(20 - CE)}.$$

which implies that $u(16) - u(CE) < \frac{16-CE}{2(20-CE)}$.

$$\frac{u(CE) - u(12)}{CE - 12} > \frac{u(CE) - u(10)}{CE - 10} = \frac{1}{2(CE - 10)},$$

which implies that $u(CE) - u(12) > \frac{CE-12}{2(CE-10)}$. Observe that for $CE > 40/3$,

$$\frac{16 - CE}{2(20 - CE)} < \frac{CE - 12}{2(CE - 10)},$$

and therefore, $u(16) - u(CE) < u(CE) - u(12)$. Since $u(CE) = 1/2$, it follows that $u(16) < 1 - u(12)$, and therefore, if $u(12) \geq \chi/2$, $u(16) < 1 - \chi/2$.

□

Proof of Proposition 3 on p. 26. First consider what happens if payoffs are negatively correlated. In any weakly undominated self-confirming equilibrium, the informed player always votes for the risky option if the informed player obtains h and the uninformed player obtains l and otherwise votes for the safe option. Suppose towards a contradiction that the uninformed player votes for the risky option with positive probability. Then his beliefs about the distribution of play on terminal nodes must be correct: the risky option is selected only when the uninformed player obtains l. Given these beliefs, his best-response is to vote for the safe option, leading to a contradiction.

Now suppose that payoffs are positively correlated. Suppose that the uninformed player believes that the informed player votes for the risky option when they both obtain l and votes for the safe option when they both obtain h. Given these beliefs, the best-response for the uninformed player is vote for the safe option. $\square$

G Instructions

Here we present the instructions that subjects saw in the HH and HC treatments. We first show the instructions of the HH treatment (where subjects play against each-other), and then the instructions of the HR treatment (where subjects play against robot players).

G.1 First treatment: subjects face other subjects

INSTRUCTIONS

This study is in 5 Parts. Only one randomly chosen Part will count for payment. In addition to what you will earn in the study, you will be paid a $7 participation fee if you complete the study.

Importantly, all Rounds and Parts of this study are independent. In other words, nothing you do in any Round or Part of this study will have any impact on your opportunities or payment in any other Round or Part of this study. In addition, since only one randomly chosen Part will be chosen for payment, it is in your best interest to treat each Part as if it was the only one that mattered for payment.

We will now hand out the instructions for Part 1 of the study. We will give you the instructions for Part 2 of the study once you have completed Part 1, for Part 3 after you have completed Part 2, etc.

Part 1

In this Part of the Study you will make decisions over the course of 9 Rounds. In each Round, you will be asked to choose between two options that determine your payoff.

Below we list exact decision problems that you will all face.

The List of Rounds in Part 1

Decision Problem	Option 1		Option 2
1	Fixed amount of \$11	Versus	Receiving \$10 or \$20 with equal chance of each.
2	Fixed amount of \$12	Versus	Receiving \$10 or \$20 with equal chance of each.
3	Fixed amount of \$13	Versus	Receiving \$10 or \$20 with equal chance of each.
4	Fixed amount of \$14	Versus	Receiving \$10 or \$20 with equal chance of each.
5	Fixed amount of \$15	Versus	Receiving \$10 or \$20 with equal chance of each.
6	Fixed amount of \$16	Versus	Receiving \$10 or \$20 with equal chance of each.
7	Fixed amount of \$17	Versus	Receiving \$10 or \$20 with equal chance of each.
8	Fixed amount of \$18	Versus	Receiving \$10 or \$20 with equal chance of each.
9	Fixed amount of \$19	Versus	Receiving \$10 or \$20 with equal chance of each.

These decision problems may appear in different order on your screen. In addition, for any given decision problem, which option appears on the left or the right of your screen may also differ from the examples above.

As you can see above, in each of the Rounds, one choice will be a fixed amount and the other will involve some uncertainty. The uncertainty can be described in the following way. The computer flips a virtual coin that lands either on heads or tails, each with an equal 50% chance. The outcome of the virtual coin flip determines your payment if you chose the uncertain option.

- if the coin lands on tails (which happens with 50% chance) you will receive $10.
- if the coin lands on heads (which happens with 50% chance) you will receive $20.

Payment: If this Part is randomly selected to count for payment in this Study, one of the 9 Rounds will be chosen to count for payment. Your earnings would be determined in the following way:

- if you chose the fixed amount, then you will earn that fixed amount;
- if you chose the option with uncertainty, your earnings depend on the result of the virtual coin flip: you receive $10 if the coin lands on tails, and you receive $20 if the coin lands on heads.

Parts 2, 3 and 4 – Preamble

Each of Parts 2, 3 and 4 consist of 2 Blocks. In each of those Parts, Block 1 consists of a series of questions that test your understanding of the instructions that are relevant to the Part you are in. In each of those Parts, Block 2 consists of several Rounds of the game itself.

In Block 2, in each Round of Parts 2, 3 and 4, you will each be randomly matched into pairs. In each Round you and the person you are matched with will be asked to vote for one of two options that determine payoffs for both you and the person you are matched with. In each Part, who you are matched with will be randomly determined at the start of each Round and nothing you do or anyone else does can influence or impact how this matching occurs. At no point will you find out with whom you were matched, nor will your actions be revealed to anyone else nor will you find out the actions of the person with whom you were matched.

If Part 2 or Part 3 or Part 4 is randomly chosen to count for payment, then you will be paid for Block 1 or Block 2 of that Part.

If Block 1 of a Part is chosen for payment, then if you answered all the questions correctly, you will earn \$10. If you make even one mistake, you will earn \$2.

If Block 2 of a Part is chosen for payment, one of the Rounds in Block 2 will be randomly selected to determine your payment.

Part 2

Part 2 is in six Rounds. As described in the preamble, in each Round you will be randomly rematched with another person in this room. In each Round you and the person you are matched with will be asked to vote for one of two options that determine payoffs for both you and the voter you are matched with. Here is an example of such a choice you can encounter in one of the Rounds (the choices you face may be different and will vary from Round to Round). Please take a moment to look at this table.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$5 Other voter's earnings: \$10	Your earnings: \$5 Other voter's earnings: \$5
vote for this option	vote for this option

Just as in the example above, each option will differ in terms of the amounts that you and/or the voter you are matched with can earn. The options also differ in how many votes are needed for that option to be the one that is selected for this Round. In each Round, one of the two options will require that both you and the voter you are matched with vote for it in order for it to be selected for this Round. The other option is selected for this Round so long as it receives at least one vote. Which option requires two votes and which option only requires at least one vote will be clearly stated before you and the voter you are matched with make your decisions.

In the example above, for the option on the right to be selected for this Round, both you and the voter you are matched with have to vote for it. On the other hand, for the option on the left to be selected for this Round, only one voter has to vote for it. In other words, if you vote for the option on the left, then it is selected for this Round regardless of what option the voter you are matched with votes for. Similarly, if the voter you are matched with votes for the option on the left it is selected for this Round regardless of which option you vote for.

Note that the option that requires one vote will always be on the left and the option that requires two votes will always be on the right hand side.

Payment: If this Part is randomly chosen to count for payment, then one Round will be randomly chosen to count for payment. In the example above, if the option on the left is selected for this Round, then you would receive \$5 and the voter you are matched with would receive \$10. If the option on the right is selected for this Round, then you would receive \$5 and the voter you are matched with would receive \$5. Which of the two options is selected for this Round will depend on what happens during the Round.

Also note that you will go through the Rounds of Part 2 without knowing what the voter you are matched with has chosen.

Do you have any questions?

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 2 is randomly chosen to count for payment, then you earn \$10 if you answer ALL the questions correctly. If you make even one mistake you will earn only \$2.

After Block 1 is over, you will play the 6 Rounds of the Part 2 game.

Part 3

Part 3 is in four Rounds. As described in the preamble, in each Round you will be randomly rematched with another person in this room. In each Round you and the person you are matched with will be asked to vote for one of two options that determine payoffs for both you and the voter you are matched with. The game in Part 3 of this study is very similar to the game you played in Part 2. The difference lies in the kinds of options you face. In this part of the study, one of the options involves uncertainty. An example of a choice with an uncertain outcome is shown below. Please take a moment to look at the table below before I describe it.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$11 Other voter's earnings: \$9	HEADS: Your earnings: \$5 Other voter's earnings: \$15 TAILS: Your earnings: \$15 Other voter's earnings: \$5
vote for this option	vote for this option

In this particular example, there is no uncertainty regarding the option on the left (the one requiring only 1 vote): if this option is the one that is selected for this Round, you would receive \$11 and the voter you are matched with would receive \$9. However, there is uncertainty regarding the option on the right (the one requiring two votes).

The uncertainty can be described in the following way. The computer throws a fair virtual coin that lands either heads or tails, each with an equal 50% chance. If it lands on heads, then the option on the right is: \$5 for you and \$15 for the voter you are matched with. If, on the other hand, it lands on tails, the option on the right is: \$15 for you and \$5 for the voter you are matched with. In other words, there is uncertainty in terms of which of the payoff pairs correspond to the option on the right: you do not know whether the payoff pair will be \$5 for you and \$15 for the voter you are matched with, or whether it will be \$15 for you and \$5 for the voter you are matched with. All you know is that the outcomes in the "uncertain" options are equally likely, each having 50% chance. In each Round the computer will flip that virtual coin before you and the voter you are matched with make your choices, but what side the coin landed on and which payoff pair that corresponds to will not be revealed to anyone.

Do you have any questions?

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 3 is randomly chosen to count for payment, then you earn \$10 if you answer ALL the questions correctly. If you make even one mistake you will earn only \$2.

After Block 1 is over, you will play the 4 Rounds of the Part 3 game.

Part 4

Part 4 is in eight Rounds. As described in the preamble, in each Round you will be randomly rematched with another person in this room. In each Round you and the person you are matched with will be asked to vote for one of two options that determine payoffs for both you and the voter you are matched with. The game in Part 4 of this study is very similar to the game you played in Part 3. The difference lies in that you OR the voter you are matched with will learn what the outcome of the virtual coin flip is before either of you vote. In other words, either you OR the voter with whom you are matched will observe whether the coin lands on heads or tails **before you have to cast a vote.** Recall the example of the previous Part in which the option requiring two votes (the one on the right) had uncertainty in terms of outcomes:

Votes needed: 1	**Votes needed: 2**
Your earnings: \$11 Other voter's earnings: \$9	HEADS: Your earnings: \$5 Other voter's earnings: \$15 TAILS: Your earnings: \$15 Other voter's earnings: \$5
vote for this option	vote for this option

Using the example above, this means that one voter in each pair will know whether the option requiring two votes leads to \$5 for you and \$15 for the other voter, or whether that option leads to \$15 for you and \$5 for the other voter.

If you are the one who learns the result of the coin flip, it means the voter you are matched with has not learned the result of the coin flip. That means before you and the other voter vote, you know exactly what happens if the option on the right is selected for this round, but the voter you are matched with does not have this information. If, on the other hand, you don't learn the result of the coin flip, it means that the voter you are matched with does know the result of the coin flip. That means before you and the other voter vote, the voter you are matched with knows exactly what happens if the option on the right is selected for this Round, but you do not have this information.

On the screen for each Round, you will know whether it is you or the voter you are matched with who has learned the result of the coin flip.

If you are the one who learns the result of the coin flip, your screen will display the relevant payoffs in black and the other payoff will be crossed out and in a lighter color. Below is an example of a Round in which you are the one who learned the result of the coin flip. In this example, you learned that the coin landed on Heads. As you can see, the payoff for Tails has been crossed out. Please take a moment to look at this example.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$11 Other voter's earnings: \$9	HEADS: Your earnings: \$5 Other voter's earnings: \$15 ~~TAILS: Your earnings: \$15~~ ~~Other voter's earnings: \$5~~
vote for this option	vote for this option

If you do not know the result of the coin flip, your screen will not have anything crossed-out. Instead, you will face a screen like the one below, where you are reminded that the voter you are matched with has learned the result of the coin flip. Your screen would look like the following:

Recall that the other voter has learned the result of the coin flip.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$11 Other voter's earnings: \$9	HEADS: Your earnings: \$5 Other voter's earnings: \$15 TAILS: Your earnings: \$15 Other voter's earnings: \$5
vote for this option	vote for this option

Do you have any questions?

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 4 is randomly chosen to count for payment, then you earn \$10 if you answer ALL the questions correctly. If you make even one mistake you will earn only \$2.

After Block 1 is over, you will play the 8 Rounds of the Part 4 game.

Part 5 Preamble

In this final Part of the study you will be assigned a Type. You will be a Type A Player or a Type B Player. Your Type will remained fixed throughout this last Part.

Each Type A Player will be randomly rematched with a Type B Player. You will not know who you are matched with. In this final Part of today's study, only Type A players make decisions that matter for payment, and these decisions affect the payoff of both the Type A Player and the Type B Player he/she is matched with.

Even though your Type will remain fixed for the rest of this study, you will not know which Type of Player you are. Since you do not know which Type of Player you are assigned to be, and since only Type A Players make decisions that matter for payment, we will ask everyone to make decisions as if they were Type A players.

Please note that your Type will remain fixed and at no point will you change roles. Your "true" Types have already been determined by the computer, and your decisions when acting as Player A CANNOT affect you or anyone else in this room if your "true" Type turns out to be Type B. In other words, if it turns out you are a Type B Player, no decision you make here can affect anyone's payoff, including your own. If it turns out your "true" Type is A, there is nothing that anyone else can do that will affect your payoff, and your decisions affect both your payoff and the payoff of the Type B Player you are matched with. Therefore, when making decisions, you should act as Player A. Further, since only "true" Type A Players make decisions that matter for payment in this study, in the remainder of the instructions we will assume you are a Type A Player.

Part 5

In this Part of the study, you will make decisions over the course of 9 Rounds.

In each Round you will have the choice between two options that determine earnings for you and the Type B player you are matched with. These choices will look like the following:

Your earnings: \$6 Type B player's earnings: \$12	Your earnings: \$11 Type B player's earnings: \$9

Here are how your payments would be determined if this were the Round that mattered for payment:

- If you chose the option on the left, then you would earn \$6 and the Type B player you are matched with would earn \$12.
- If instead you chose the option on the right, then you would earn \$11 and the Type B player you are matched with would earn \$9.

In each Round, you are the one whose decision will matter. That is, it is your choice of option that will be selected for each Round.

Remember that you will not change roles. So as a Type A Player, your payoff will never be determined by someone else in this room. Also remember that only one Part of the study will be chosen to count for payment. If this Part is chosen to count, only *one* Round will matter for payment. So it is in your best interest to treat each Round as if it were the one that mattered for payment.

Do you have any questions?

G.2 HR treatment: subjects face robot players

INSTRUCTIONS

This study is in 5 Parts. Only one randomly chosen Part will count for payment. In addition to what you will earn in the study, you will be paid a $7 participation fee if you complete the study.

Importantly, all Rounds and Parts of this study are independent. In other words, nothing you do in any Round or Part of this study will have any impact on your opportunities or payment in any other Round or Part of this study. In addition, since only one randomly chosen Part will be chosen for payment, it is in your best interest to treat each Part as if it was the only one that mattered for payment.

We will now hand out the instructions for Part 1 of the study. We will give you the instructions for Part 2 of the study once you have completed Part 1, for Part 3 after you have completed Part 2, etc.

Part 1

In this Part of the Study you will make decisions over the course of 9 Rounds. In each Round, you will be asked to choose between two options that determine your payoff.

Below we list exact decision problems that you will all face.

The List of Rounds in Part 1

Decision Problem	**Option 1**		**Option 2**
1	Fixed amount of $11	Versus	Receiving $10 or $20 with equal chance of each.
2	Fixed amount of $12	Versus	Receiving $10 or $20 with equal chance of each.
3	Fixed amount of $13	Versus	Receiving $10 or $20 with equal chance of each.
4	Fixed amount of $14	Versus	Receiving $10 or $20 with equal chance of each.
5	Fixed amount of $15	Versus	Receiving $10 or $20 with equal chance of each.
6	Fixed amount of $16	Versus	Receiving $10 or $20 with equal chance of each.
7	Fixed amount of $17	Versus	Receiving $10 or $20 with equal chance of each.
8	Fixed amount of $18	Versus	Receiving $10 or $20 with equal chance of each.
9	Fixed amount of $19	Versus	Receiving $10 or $20 with equal chance of each.

These decision problems may appear in different order on your screen. In addition, for any given decision problem, which option appears on the left or the right of your screen may also differ from the examples above.

As you can see above, in each of the Rounds, one choice will be a fixed amount and the other will involve some uncertainty. The uncertainty can be described in the following way. A virtual coin is flipped, that lands either on heads or tails, each with an equal 50% chance. The outcome of the virtual coin flip determines your payment if you chose the uncertain option.

- if the coin lands on tails (which happens with 50% chance) you will receive $10.
- if the coin lands on heads (which happens with 50% chance) you will receive $20.

Payment: If this Part is randomly selected to count for payment in this Study, one of the 9 Rounds will be chosen to count for payment. Your earnings would be determined in the following way:

- if you chose the fixed amount, then you will earn that fixed amount;
- if you chose the option with uncertainty, your earnings depend on the result of the virtual coin flip: you receive $10 if the coin lands on tails, and you receive $20 if the coin lands on heads.

Parts 2, 3 and 4 – Preamble

Each of Parts 2, 3 and 4 consist of 2 Blocks. In each of those Parts, Block 1 consists of a series of questions that test your understanding of the instructions that are relevant to the Part you are in. In each of those Parts, Block 2 consists of several Rounds of the game itself.

In Block 2, in each Round of Parts 2, 3 and 4, you will each be matched with a computer player. In each Round you and the computer player you are matched with will vote for one of two options that determine your payoff. How the computer player has been programmed to vote will be described to you in each Part before you cast your vote. How the computer player has been programmed to vote will vary from Part to Part.

If Part 2 or Part 3 or Part 4 is randomly chosen to count for payment, then you will be paid for Block 1 or Block 2 of that Part.

If Block 1 of a Part is chosen for payment, then if you answered all the questions correctly, you will earn $10. If you make even one mistake, you will earn $2.

If Block 2 of a Part is chosen for payment, one of the Rounds in Block 2 will be randomly selected to determine your payment.

Part 2

Part 2 is in six Rounds. As described in the preamble, in each Round you will be matched with a computer player. In each Round you will be asked to vote for one of two options that determine how much you will earn. It will also determine how many virtual (imaginary) dollars the computer will earn.

Note that how many virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever. We will describe how the computer player is programmed to cast its vote in this Part of the Study, only after describing the kinds of choices you will face.

We will describe the kinds of choices you will face by using an example. Below is an example of a choice you can encounter in one of the Rounds (the choices you face may be different and will vary from Round to Round). Please take a moment to look at this table.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$5 Computer player: \$10 virtual (imaginary) dollars	Your earnings: \$6 Computer player: \$5 virtual (imaginary) dollars
vote for this option	vote for this option

Just as in the example above, each option will differ in terms of the amount that you and/or the computer player can earn. The options also differ in how many votes are needed for that option to be the one that is selected for this Round. In each Round, one of the two options will require that both you and the computer player you are matched with vote for it in order for it to be selected for this Round. The other option is selected for this Round so long as it receives at least one vote. Which option requires two votes and which option only requires at least one vote will be clearly stated before you make your decisions.

In the example above, for the option on the right to be selected for this Round, both you and the computer player you are matched with have to vote for it. On the other hand, for the option on the left to be selected for this Round, only one voter has to vote for it. In other words, if you vote for the option on the left, then it is selected for this Round regardless of what option the computer player you are matched with votes for. Similarly, if the computer player you are matched with votes for the option on the left it is selected for this Round regardless of which option you vote for.

Note that the option that requires one vote will always be on the left and the option that requires two votes will always be on the right hand side.

Payment: If this Part is randomly chosen to count for payment, then one Round will be randomly chosen to count for payment. In the example above, if the option on the right is selected for this Round, then you would receive \$6 and the computer player would receive

5 virtual (imaginary) dollars. If the option on the left is selected for this Round, then you would receive $5 and the computer player would receive 10 virtual (imaginary) dollars. Which of the two options is selected for this Round will depend on what happens during the Round.

Recall that the virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever.

Before telling you how the computer player has been programmed to vote, we will ask you 6 questions that test your understanding of these instructions. After you have answered these 6 questions, we will tell you how the computer player has been programmed to vote in Part 2 and then ask you one additional understanding question.

Do you have any questions?

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 2 is randomly chosen to count for payment, then you earn $10 if you answer ALL the questions correctly. If you make even one mistake you will earn only $2.

[*Note: subjects received what follows after the Part 2 Block 1 questions.*]

In Part 2 the computer player has been programmed to **ALWAYS choose the option that gives it the highest number of virtual (imaginary) dollars.** That is, the computer player will look at which option gives it the highest number of virtual (imaginary) dollars, and will vote for that one.

Additional understanding question:

1. Which option will the computer player vote for? [the one that requires 1 vote only; the one that requires 2 votes; it will randomly choose which option to vote for, each option having an equal chance; it will vote for the option that gives it the highest amount of virtual (imaginary) dollars.]

Part 3

Part 3 is in four Rounds. As described in the preamble, in each Round you will be matched with a computer player. In each Round you will be asked to vote for one of two options that determine how much you will earn. It will also determine how many virtual (imaginary) dollars the computer will earn. Note that how many virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever.

In this Part of the Study, the way the computer player has been programmed to vote is different than in Part 2. Before we describe how the computer player is programmed to cast its vote in this Part of the study, we'll start by showing you an example of what you might see in a Round.

The game in Part 3 of this study is very similar to the game you played in Part 2. The difference lies in the kinds of options you face. In this part of the study, one of the options involves uncertainty. An example of a choice with an uncertain outcome is shown below. Please take a moment to look at the table below before I describe it.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$11 Computer player: \$9 virtual (imaginary) dollars	HEADS: Your earnings: \$5 Computer player: \$15 virtual (imaginary) dollars TAILS: Your earnings: \$15 Computer player: \$5 virtual (imaginary) dollars
vote for this option	vote for this option

In this particular example, there is no uncertainty regarding the option on the left (the one requiring only 1 vote): if this option is the one that is selected for this Round, you would receive \$11 and the computer player you are matched with would receive \$9 virtual (imaginary) dollars. However, there is uncertainty regarding the option on the right (the one requiring two votes).

The uncertainty can be described in the following way. As before, a coin is flipped and lands either heads or tails, each with an equal 50% chance. If it lands on heads, then the option on the right is: \$5 for you and \$15 virtual (imaginary) dollars for the computer player you are matched with. If, on the other hand, it lands on tails, the option on the right is: \$15 for you and \$5 virtual (imaginary) dollars for the computer player you are matched with. In other words, there is uncertainty in terms of which of the payoff pairs correspond to the option on the right: you do not know whether the payoff pair will be \$5 for you and \$15 virtual (imaginary) dollars for the computer player you are matched with, or whether it will be \$15 for you and \$5 virtual (imaginary) dollars for the computer player you are matched with. All you know is that the outcomes in the "uncertain" options are equally likely, each having 50% chance. In each Round the coin will be flipped before you and the computer player you are matched with make your choices, but what side the coin landed on and which payoff pair that corresponds to will not be revealed to anyone.

Recall that the virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever.

Do you have any questions?

Before telling you how the computer player has been programmed to vote, we will ask you 4 questions that test your understanding of these instructions. After you have answered these 4 questions, we will tell you how the computer player has been programmed to vote in Part 3, and then ask you one additional understanding question.

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 3 is randomly chosen to count for payment, then you earn $10 if you answer ALL the questions correctly. If you make even one mistake you will earn only $2.

After Block 1 is over, you will play the 4 Rounds of the Part 3 game.

[*Note: subjects received what follows after the Part 3 Block 1 questions.*]

In Part 3 the computer player has been programmed to **ALWAYS choose the option on the right**. That is, it will vote for the option on the right (the one requiring 2 votes) no matter what.

Part 4

Part 4 is in four Rounds. As described in the preamble, in each Round you will be matched with a computer player. In each Round you will be asked to vote for one of two options that determine how much you will earn. It will also determine how many virtual (imaginary) dollars the computer will earn. Note that how many virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever.

The game in Part 4 of this study is very similar to the game you played in Part 3 except for two things:

1. The first difference lies in that the computer player you are matched with will learn what the outcome of the coin flip is before it votes. You, however, will not know what the outcome of the coin flip is and will have to cast your vote without knowing the outcome of the coin flip.
2. The second is that the computer player has been programmed to vote for the option that gives it the highest number of virtual (imaginary) dollars. That is, after learning the outcome of the coin flip, it will look at which option gives it the highest amount of virtual (imaginary) dollars and vote for that one.

The kind of screen you face will look very similar to the kind of screen you faced in Part 3, except that you will be reminded that the virtual player knows the outcome of the coin flip AND always votes for the option that gives it the highest amount of virtual (imaginary) dollars:

Recall that the computer player has learned the result of the coin flip and always votes for the option that gives it the highest amount of virtual (imaginary) dollars.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$11 Computer player: \$9 virtual (imaginary) dollars	HEADS: Your earnings: \$5 Computer player: \$15 virtual (imaginary) dollars TAILS: Your earnings: \$15 Computer player: \$5 virtual (imaginary) dollars
vote for this option	vote for this option

Recall that the virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever.

Do you have any questions?

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 4 is randomly chosen to count for payment, then you earn \$10 if you answer ALL the questions correctly. If you make even one mistake, you will earn only \$2.

After Block 1 is over, you will play the four Rounds of the Part 4 game.

Part 5

Part 5 also consists of two Blocks. In Part 5, Block 1 consists of 4 questions about the game you just played in Part 4. These questions will appear on a series of separate screens. If Block 1 of Part 5 is chosen for payment, you will be paid \$10 if you answer all these questions correctly. If you make even one mistake, you will only earn \$2.

Recall that in Part 4, in each Round, the computer player learned the outcome of the coin flip before it voted, and then always voted for the option that gave it the highest amount of virtual (imaginary) dollars.

After Block 1 is over, we will hand out instructions for Block 2.

Block 2:

You now are going to play 4 additional Rounds, just like the ones you played in Part 4. Recall that in those Rounds, the computer player you are matched with ALWAYS knows the result of the coin flip before it votes, and then ALWAYS votes for the option that will give it the highest number of of virtual (imaginary) dollars.

G.3 Full Feedback treatment: subjects face robot players and receive partial feedback in Part 4 over 20 rounds

.

The instructions are identical to the HR treatment except for what happens in Part 4. Below we present the Part 4 instructions for this treatment.

Part 4

Part 4 is in 20 Rounds. As described in the preamble, in each Round you will be matched with a computer player. In each Round you will be asked to vote for one of two options that determine how much you will earn. It will also determine how many virtual (imaginary) dollars the computer will earn.

Note that how many virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever.

The game in Part 4 of this study is very similar to the game you played in Part 3 except for three things:

1. The first difference lies in that the computer player you are matched with will learn what the outcome of the coin flip is before it votes. You, however, will not know what the outcome of the coin flip is and will have to cast your vote without knowing the outcome of the coin flip.

2. The second is that the computer player has been programmed to vote for the option that gives it the highest number of virtual (imaginary) dollars. That is, after learning the outcome of the coin flip, it will look at which option gives it the highest amount of virtual (imaginary) dollars and vote for that one.

3. The third is that in each round, after you have cast your vote, you will receive information on what the outcome of the virtual coin flip was, and you will be told which option the computer player voted for. In addition, you will be told which option was selected for each round and you will be told how much you would earn if that were the round randomly chosen for payment.

The kind of screen you face will look very similar to the kind of screen you faced in Part 3, except that you will be reminded that the virtual player knows the outcome of the coin flip AND always votes for the option that gives it the highest amount of virtual (imaginary) dollars:

Recall that the computer player has learned the result of the coin flip and always votes for the option that gives it the highest amount of virtual (imaginary) dollars.

Votes needed: 1	**Votes needed: 2**
Your earnings: \$11 Computer player: \$9 virtual (imaginary) dollars	HEADS: Your earnings: \$5 Computer player: \$15 virtual (imaginary) dollars TAILS: Your earnings: \$15 Computer player: \$5 virtual (imaginary) dollars
vote for this option	vote for this option

Recall that the virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever. So, the only thing that determines your earnings in this study are the number of dollars you earn in the option that is selected for a Round.

Do you have any questions?

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 4 is randomly chosen to count for payment, then you earn \$10 if you answer ALL the questions correctly. If you make even one mistake, you will earn only \$2.

After Block 1 is over, you will play the 20 Rounds of the Part 4 game.

G.3.1 Screen Shot - Example of Feedback Screen

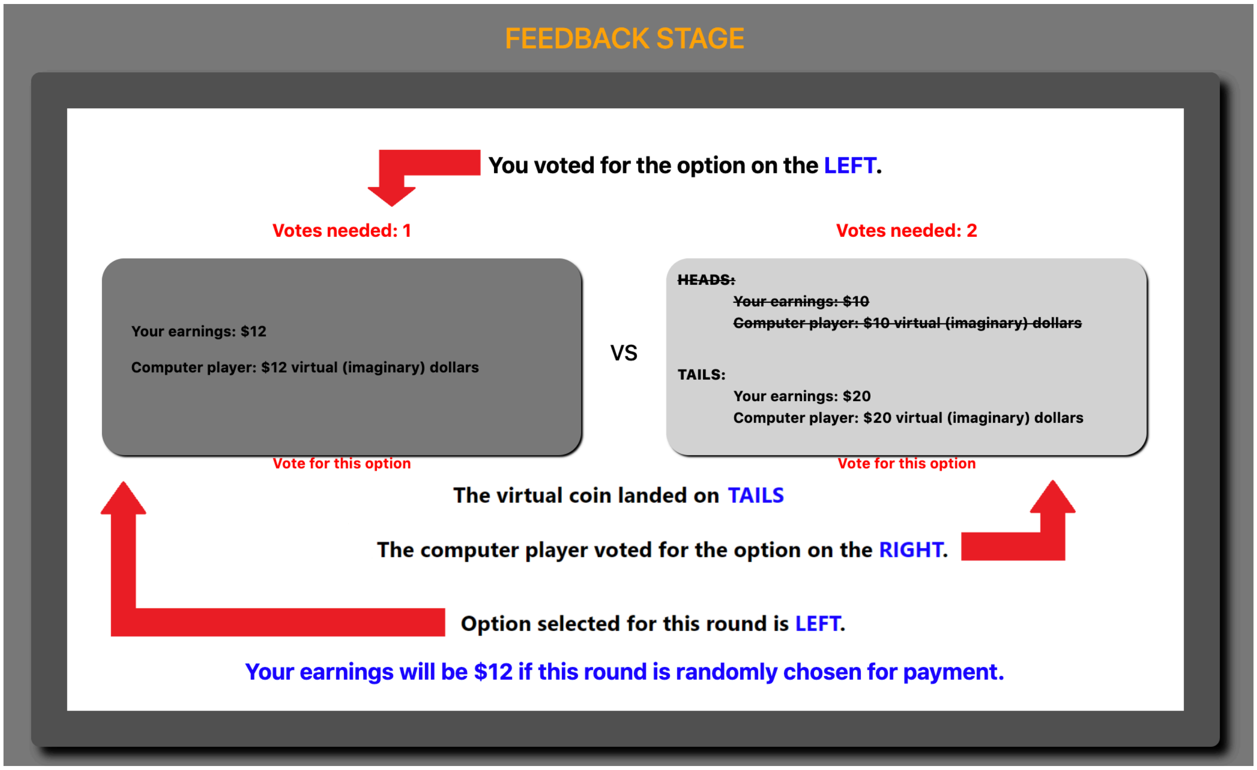


Figure 3: Full Feedback Example.

G.4 Partial Feedback treatment: subjects face robot players and receive partial feedback in Part 4 over 20 rounds

.

The instructions are identical to the HR treatment except for what happens in Part 4. Below we present the Part 4 instructions for this treatment.

Part 4

Part 4 is in 20 Rounds. As described in the preamble, in each Round you will be matched with a computer player. In each Round you will be asked to vote for one of two options that determine how much you will earn. It will also determine how many virtual (imaginary) dollars the computer will earn.

Note that how many virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever.

The game in Part 4 of this study is very similar to the game you played in Part 3 except for three things:

1. The first difference lies in that the computer player you are matched with will learn what the outcome of the coin flip is before it votes. You, however, will not know what the outcome of the coin flip is and will have to cast your vote without knowing the outcome of the coin flip.

2. The second is that the computer player has been programmed to vote for the option that gives it the highest number of virtual (imaginary) dollars. That is, after learning the outcome of the coin flip, it will look at which option gives it the highest amount of virtual (imaginary) dollars and vote for that one.

3. The third is that in each round, after you have cast your vote, you will be told how much you would earn if that were the round randomly chosen for payment.

The kind of screen you face will look very similar to the kind of screen you faced in Part 3, except that you will be reminded that the virtual player knows the outcome of the coin flip AND always votes for the option that gives it the highest amount of virtual (imaginary) dollars:

Recall that the computer player has learned the result of the coin flip and always votes for the option that gives it the highest amount of virtual (imaginary) dollars.

Votes needed: 1	**Votes needed: 2**
Your earnings: $11 Computer player: $9 virtual (imaginary) dollars	HEADS: Your earnings: $5 Computer player: $15 virtual (imaginary) dollars TAILS: Your earnings: $15 Computer player: $5 virtual (imaginary) dollars
vote for this option	vote for this option

Recall that the virtual (imaginary) dollars the computer player earns will have no impact on you or anyone else at any point, ever. So, the only thing that determines your earnings in this study are the number of dollars you earn in the option that is selected for a Round.

Do you have any questions?

We will now begin Block 1 in which you will be asked questions that test your understanding of this game. If Block 1 of Part 4 is randomly chosen to count for payment, then you earn $10 if you answer ALL the questions correctly. If you make even one mistake, you will earn only $2.

After Block 1 is over, you will play the 20 Rounds of the Part 4 game.

G.4.1 Screen Shot - Example of Feedback Screen

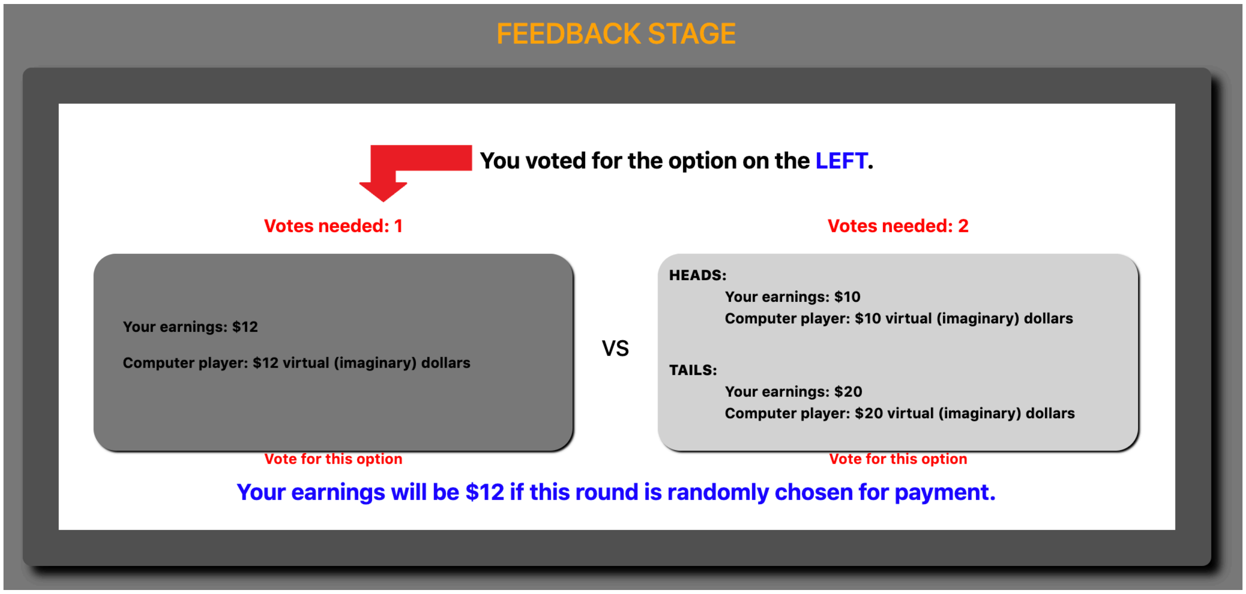


Figure 4: Partial Feedback Example.

H Understanding Questions

Below we present screen shots for each of our understanding questions.

H.1 Treatment 1: subjects play against each other

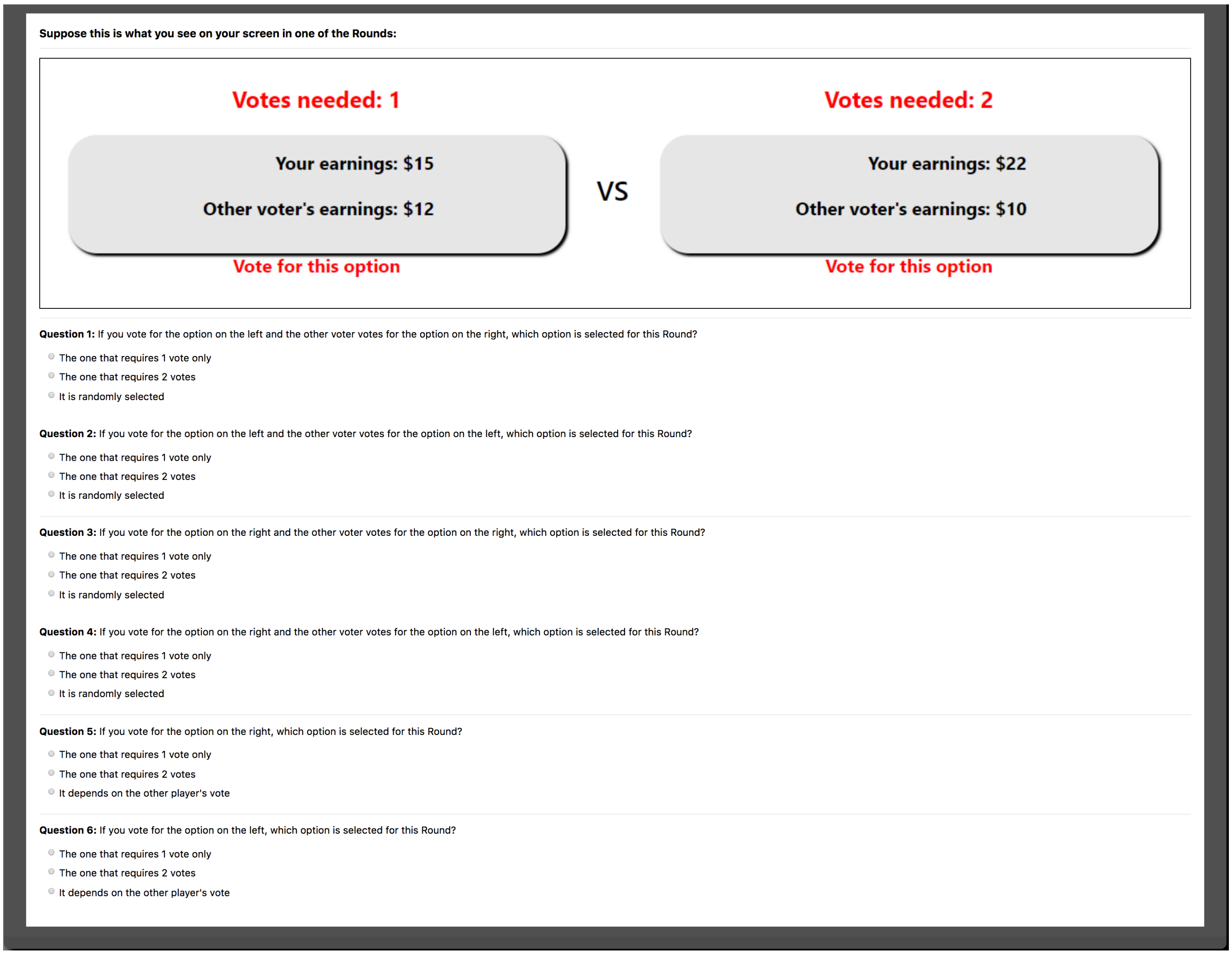
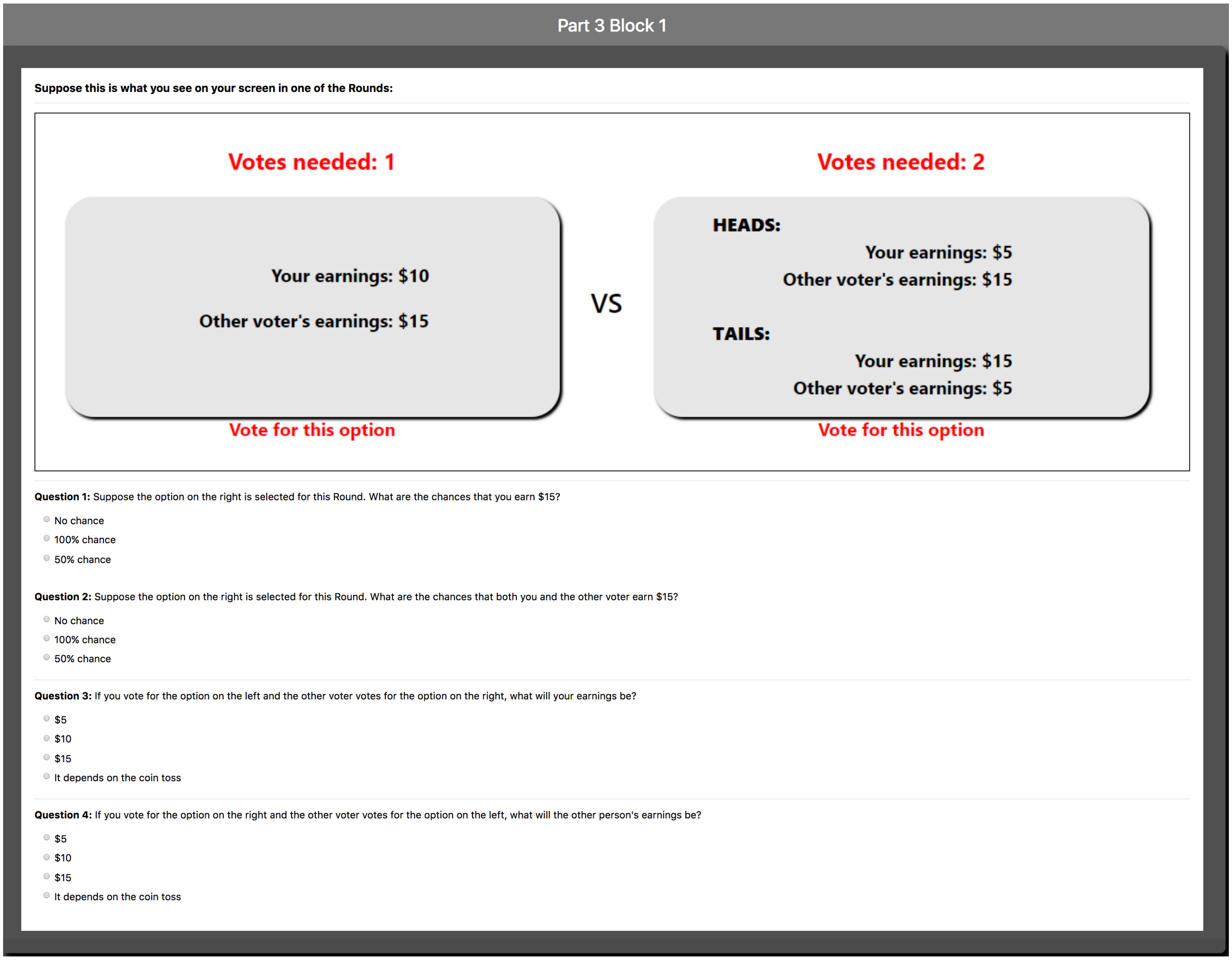


Figure 5: Part 2 Understanding questions.



Part 3 Block 1

Suppose this is what you see on your screen in one of the Rounds:

Votes needed: 1		Votes needed: 2
Your earnings: $10 Other voter's earnings: $15	VS	HEADS: Your earnings: $5 Other voter's earnings: $15 TAILS: Your earnings: $15 Other voter's earnings: $5
Vote for this option		Vote for this option

Question 1: Suppose the option on the right is selected for this Round. What are the chances that you earn $15?

- No chance
- 100% chance
- 50% chance

Question 2: Suppose the option on the right is selected for this Round. What are the chances that both you and the other voter earn $15?

- No chance
- 100% chance
- 50% chance

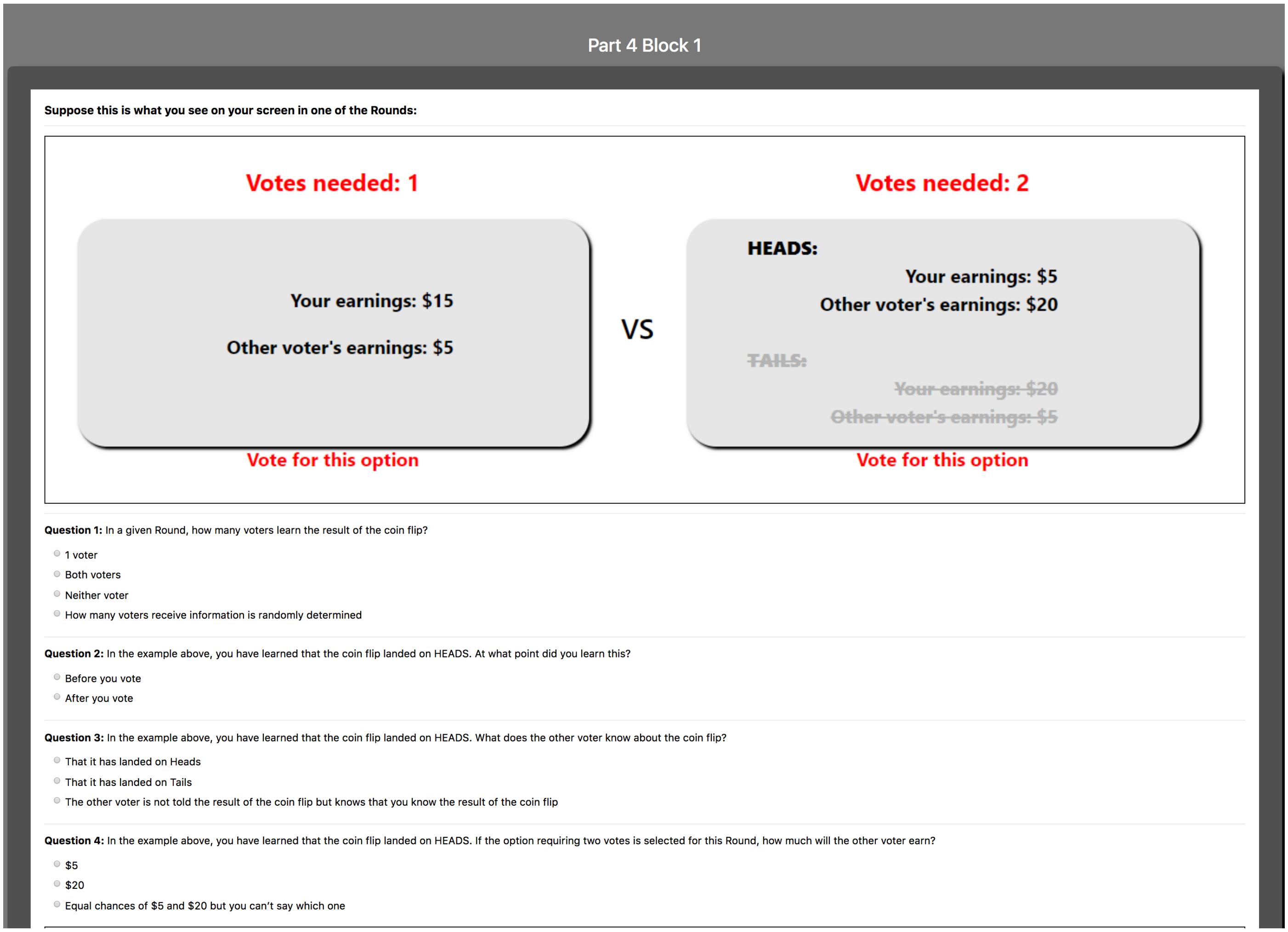
Question 3: If you vote for the option on the left and the other voter votes for the option on the right, what will your earnings be?

- $5
- $10
- $15
- It depends on the coin toss

Question 4: If you vote for the option on the right and the other voter votes for the option on the left, what will the other person's earnings be?

- $5
- $10
- $15
- It depends on the coin toss

Figure 6: Part 3 Understanding questions.



Part 4 Block 1

Suppose this is what you see on your screen in one of the Rounds:

Votes needed: 1

Your earnings: $15

Other voter's earnings: $5

Vote for this option

VS

Votes needed: 2

HEADS:

Your earnings: $5

Other voter's earnings: $20

~~TAILS:~~

~~Your earnings: $20~~

~~Other voter's earnings: $5~~

Vote for this option

Question 1: In a given Round, how many voters learn the result of the coin flip?

- 1 voter
- Both voters
- Neither voter
- How many voters receive information is randomly determined

Question 2: In the example above, you have learned that the coin flip landed on HEADS. At what point did you learn this?

- Before you vote
- After you vote

Question 3: In the example above, you have learned that the coin flip landed on HEADS. What does the other voter know about the coin flip?

- That it has landed on Heads
- That it has landed on Tails
- The other voter is not told the result of the coin flip but knows that you know the result of the coin flip

Question 4: In the example above, you have learned that the coin flip landed on HEADS. If the option requiring two votes is selected for this Round, how much will the other voter earn?

- $5
- $20
- Equal chances of $5 and $20 but you can't say which one

Figure 7: Part 4 Understanding questions (questions 1-4).

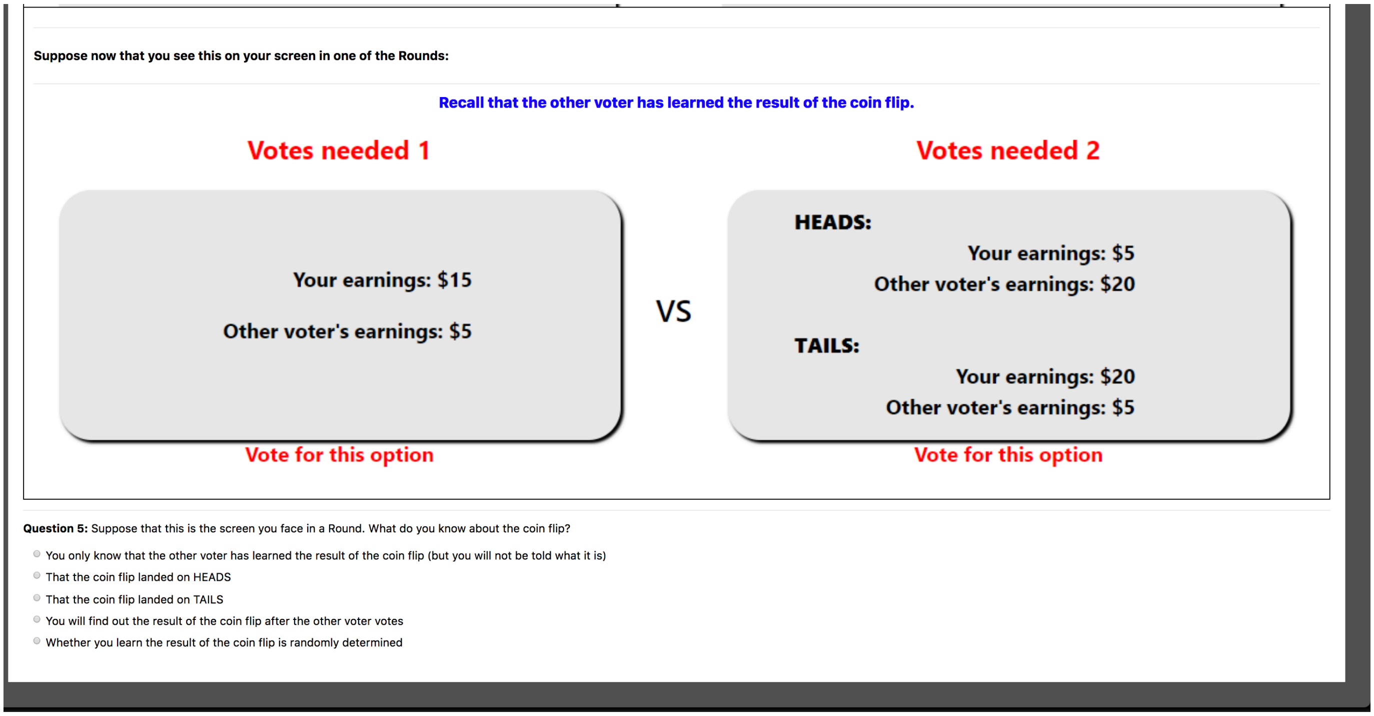


Figure 8: Part 4 Understanding questions (question 5).

H.2 Treatment 2: subjects play against a computerized robot player

Suppose this is what you see on your screen in one of the Rounds:

Votes needed: 1		Votes needed: 2
Your earnings: $15 Computer player: $12 virtual (imaginary) dollars	VS	Your earnings: $22 Computer player: $10 virtual (imaginary) dollars
Vote for this option		Vote for this option

Question 1: If you vote for the option on the left and the computer player votes for the option on the right, which option is selected for this Round?

- The one that requires 1 vote only
- The one that requires 2 votes
- It is randomly selected

Question 2: If you vote for the option on the left and the computer player votes for the option on the left, which option is selected for this Round?

- The one that requires 1 vote only
- The one that requires 2 votes
- It is randomly selected

Question 3: If you vote for the option on the right and the computer player votes for the option on the right, which option is selected for this Round?

- The one that requires 1 vote only
- The one that requires 2 votes
- It is randomly selected

Figure 9: Part 2 Understanding questions (questions 1-3).

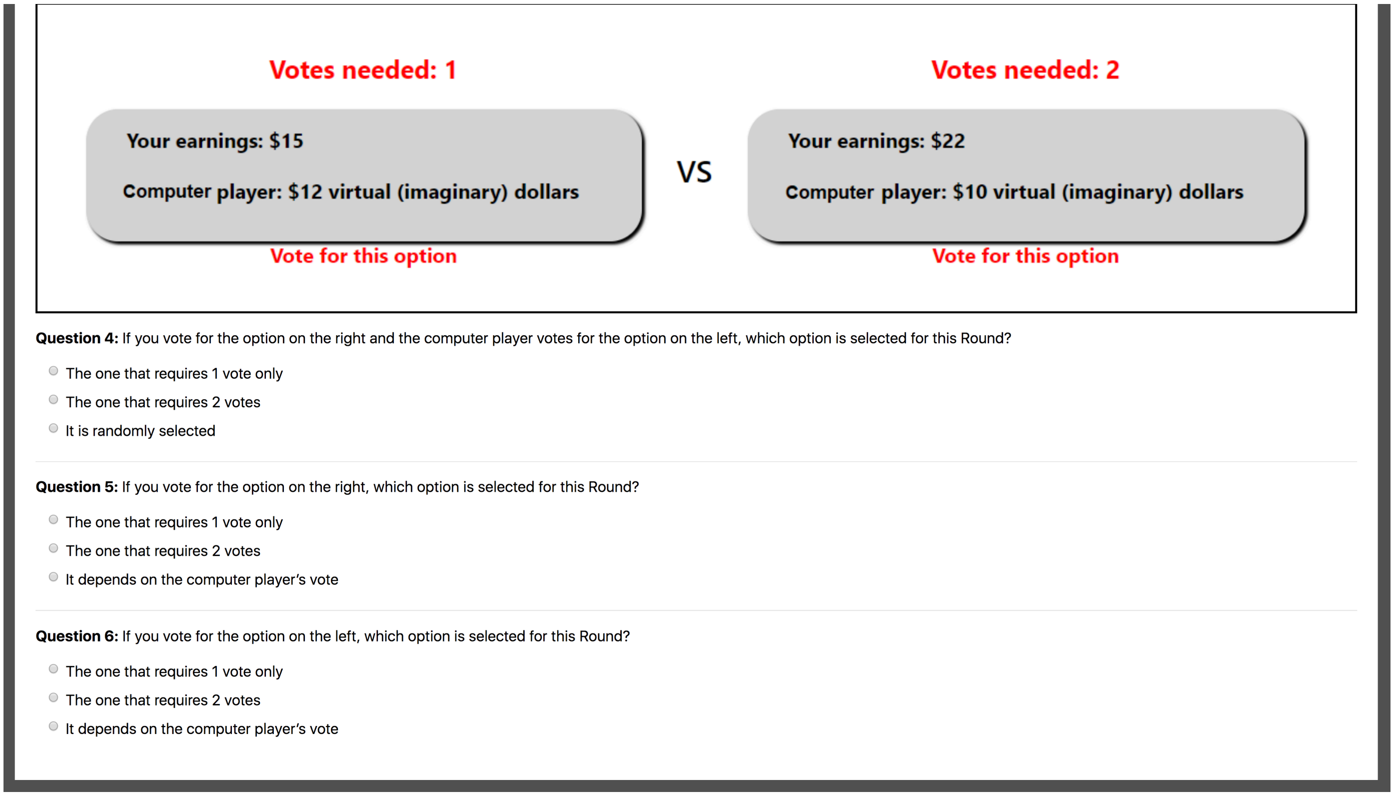


Figure 10: Part 2 Understanding question (questions 4-6).

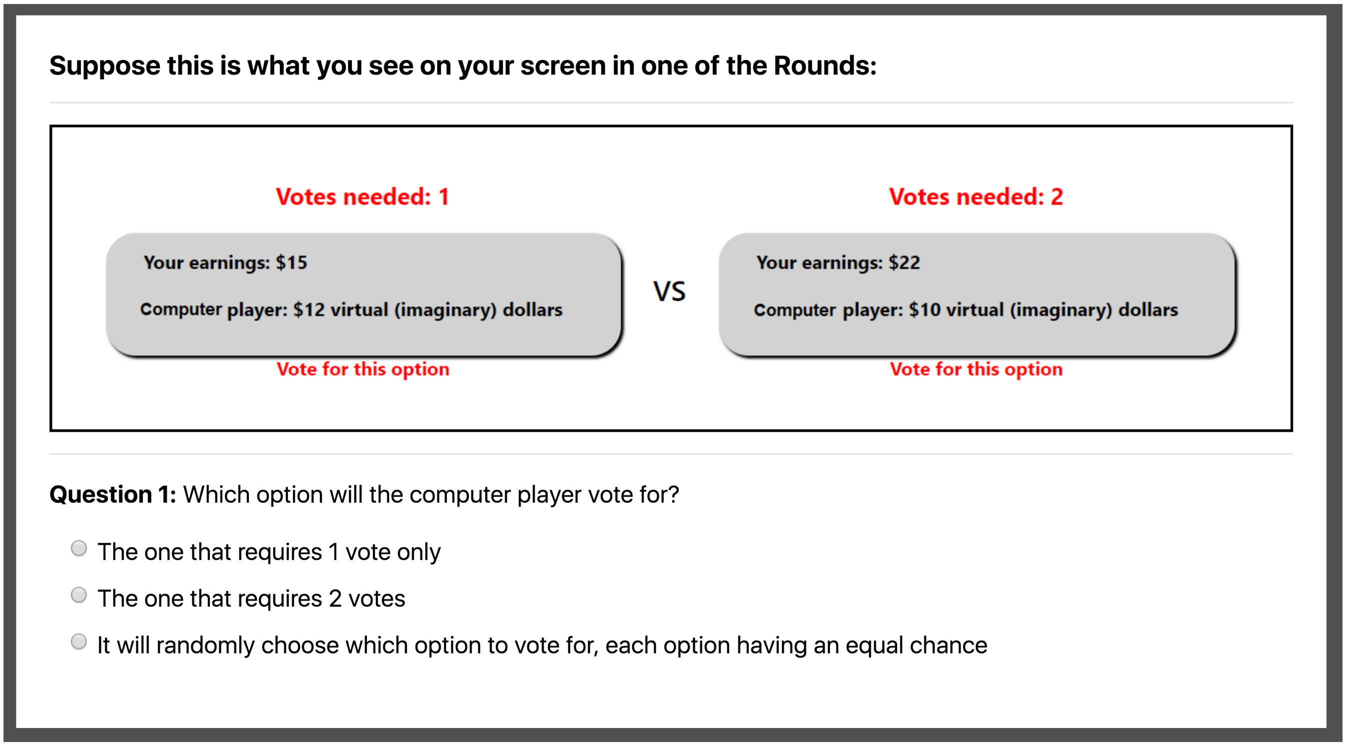


Figure 11: Part 2 Understanding question (question 7).

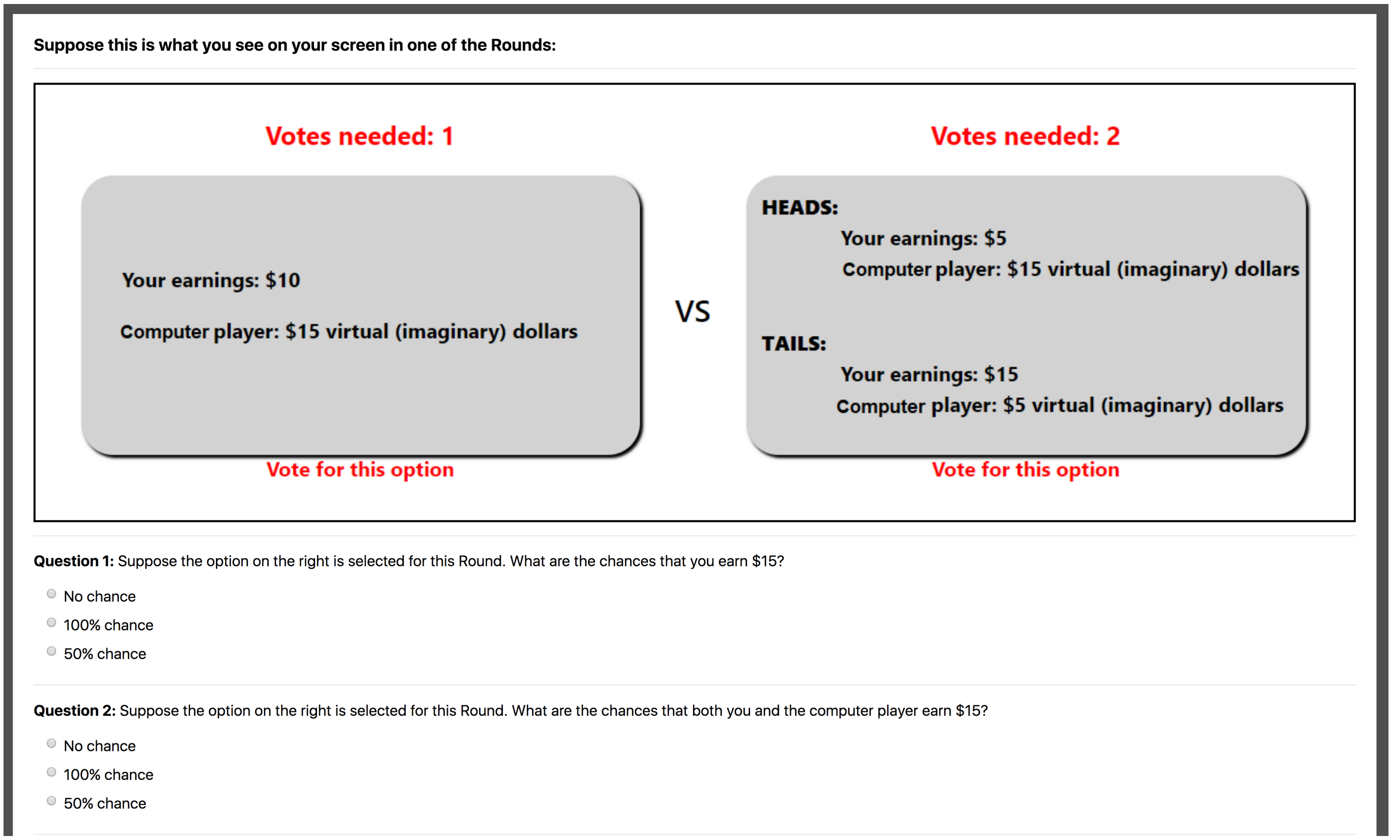


Figure 12: Part 3 Understanding questions (questions 1-2).

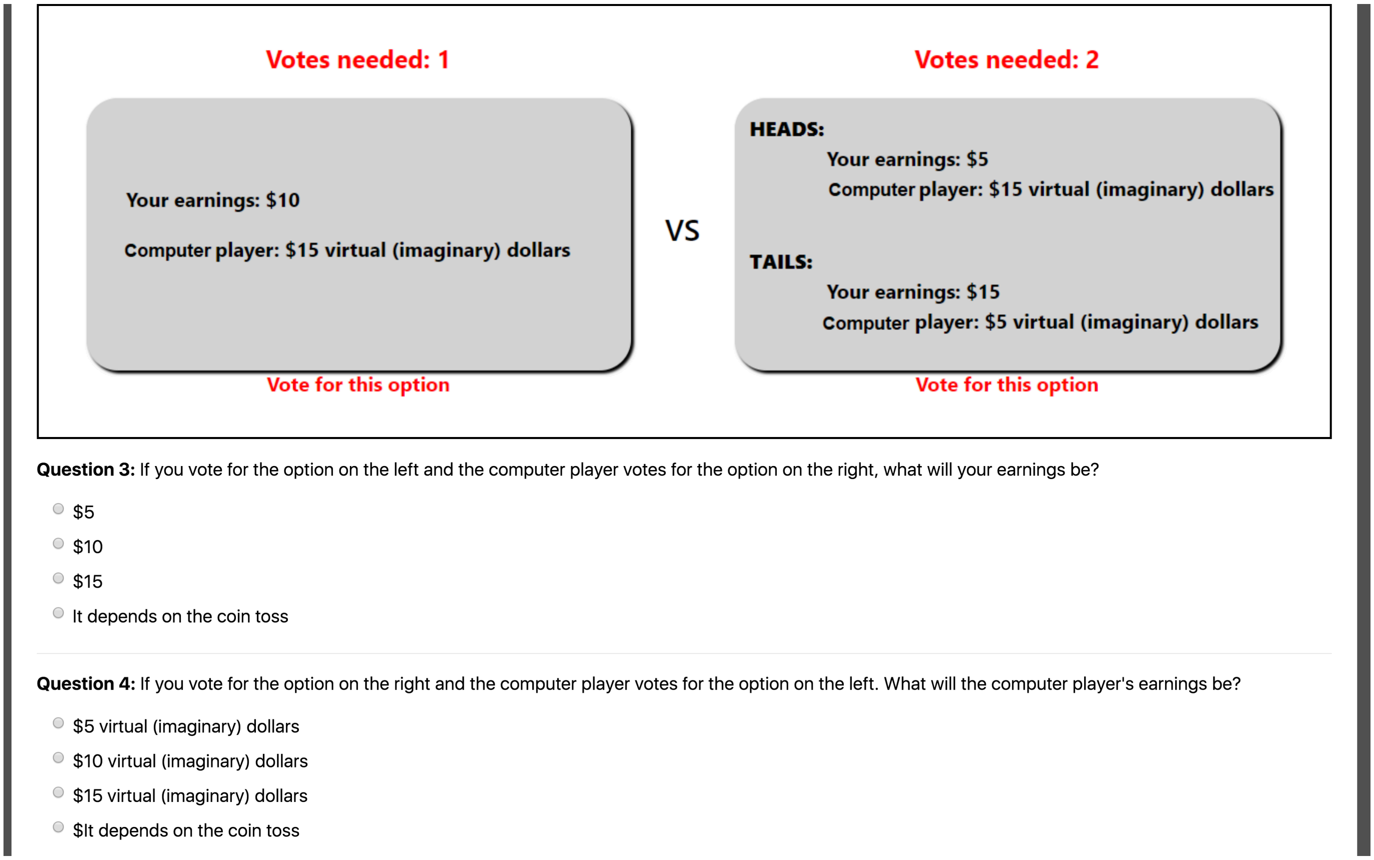


Figure 13: Part 3 Understanding questions (questions 3-4).

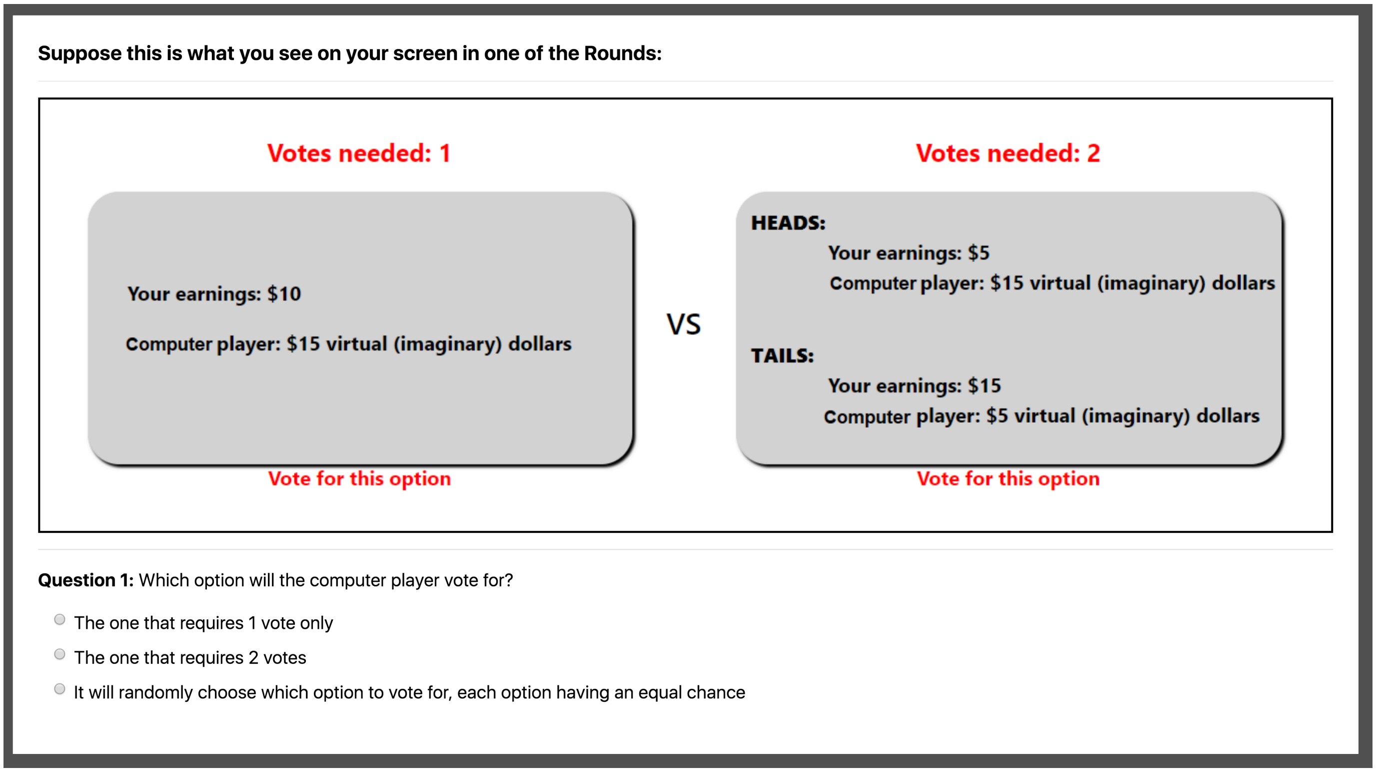
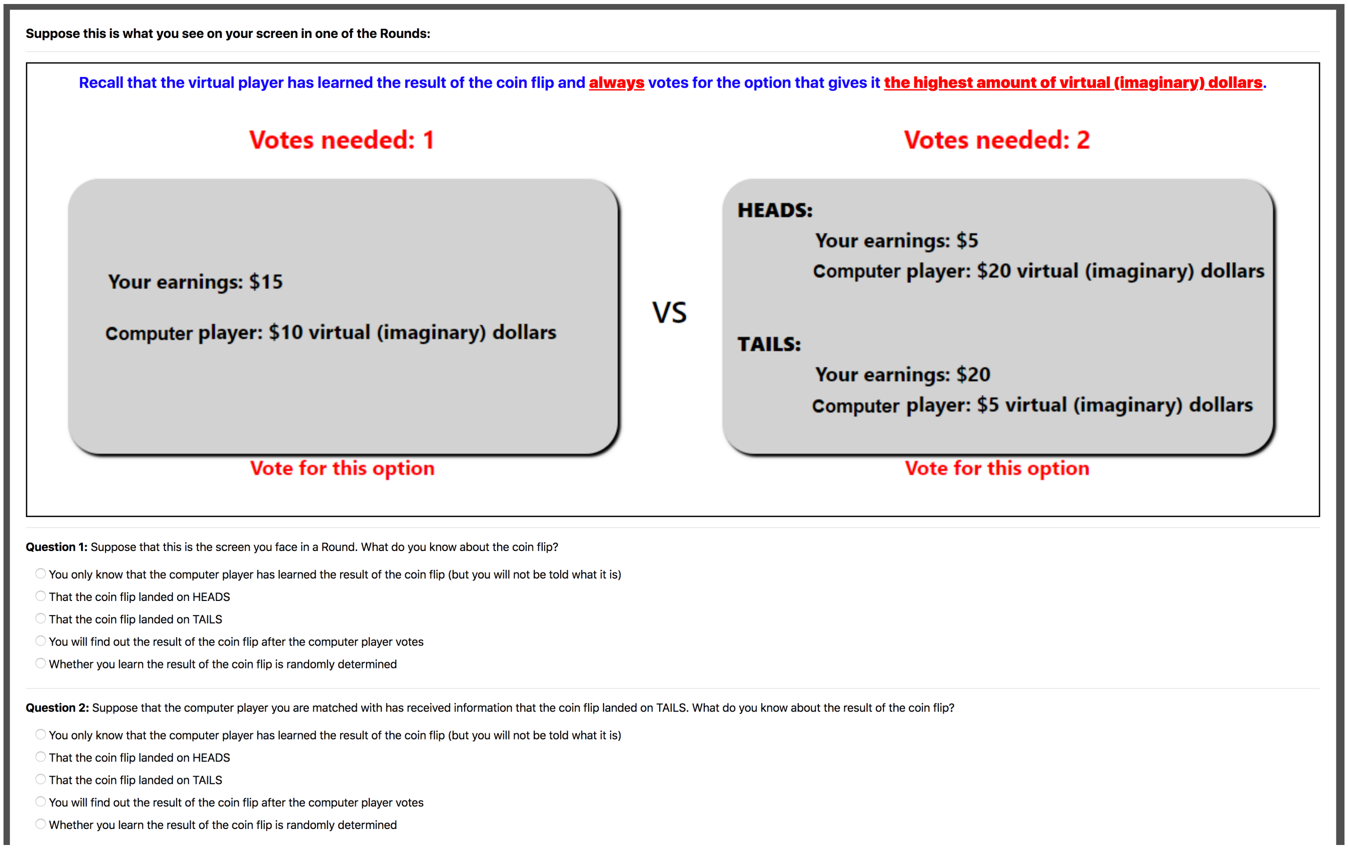


Figure 14: Part 3 Understanding questions (question 5).



Suppose this is what you see on your screen in one of the Rounds:

Recall that the virtual player has learned the result of the coin flip and **always** votes for the option that gives it **the highest amount of virtual (imaginary) dollars**.

Votes needed: 1

Your earnings: $15

Computer player: $10 virtual (imaginary) dollars

Vote for this option

VS

Votes needed: 2

HEADS:

Your earnings: $5

Computer player: $20 virtual (imaginary) dollars

TAILS:

Your earnings: $20

Computer player: $5 virtual (imaginary) dollars

Vote for this option

Question 1: Suppose that this is the screen you face in a Round. What do you know about the coin flip?

- You only know that the computer player has learned the result of the coin flip (but you will not be told what it is)
- That the coin flip landed on HEADS
- That the coin flip landed on TAILS
- You will find out the result of the coin flip after the computer player votes
- Whether you learn the result of the coin flip is randomly determined

Question 2: Suppose that the computer player you are matched with has received information that the coin flip landed on TAILS. What do you know about the result of the coin flip?

- You only know that the computer player has learned the result of the coin flip (but you will not be told what it is)
- That the coin flip landed on HEADS
- That the coin flip landed on TAILS
- You will find out the result of the coin flip after the computer player votes
- Whether you learn the result of the coin flip is randomly determined

Figure 15: Part 4 Understanding questions (questions 1-2).

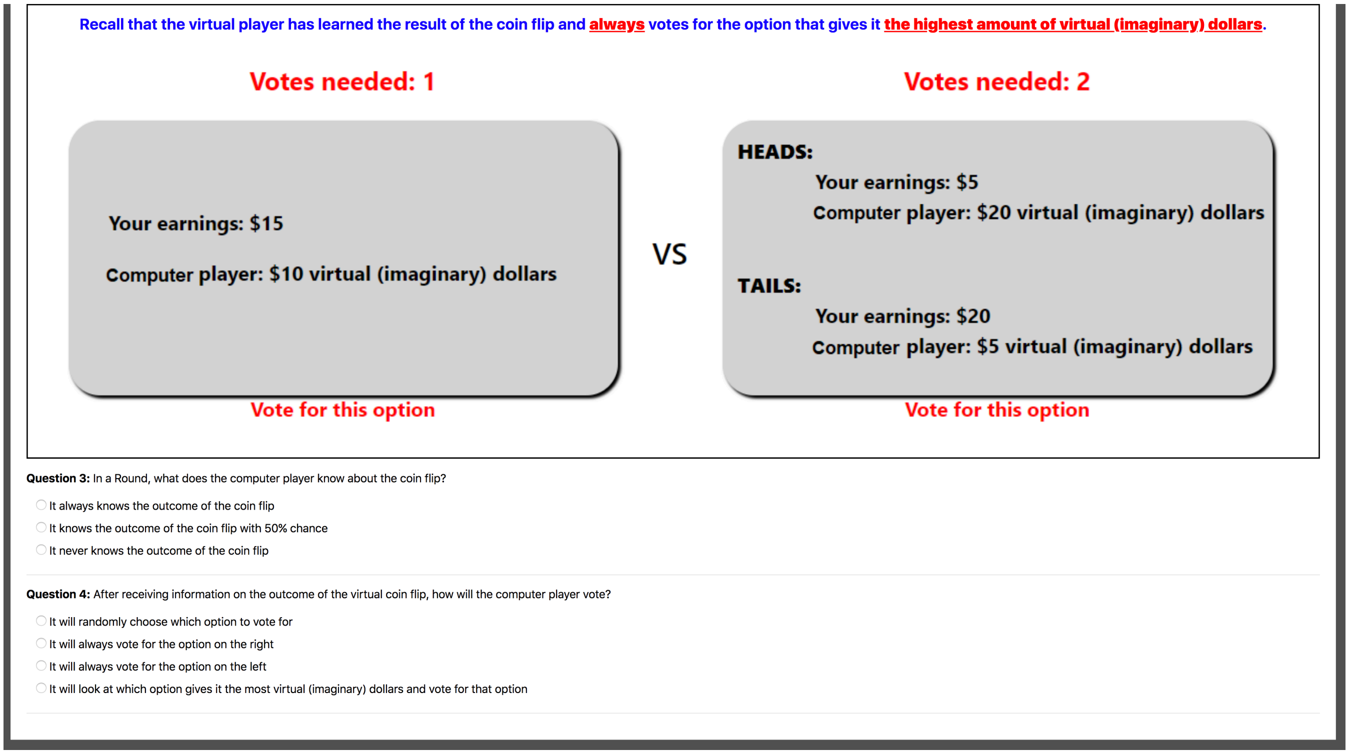


Figure 16: Part 4 Understanding questions (questions 3-4).

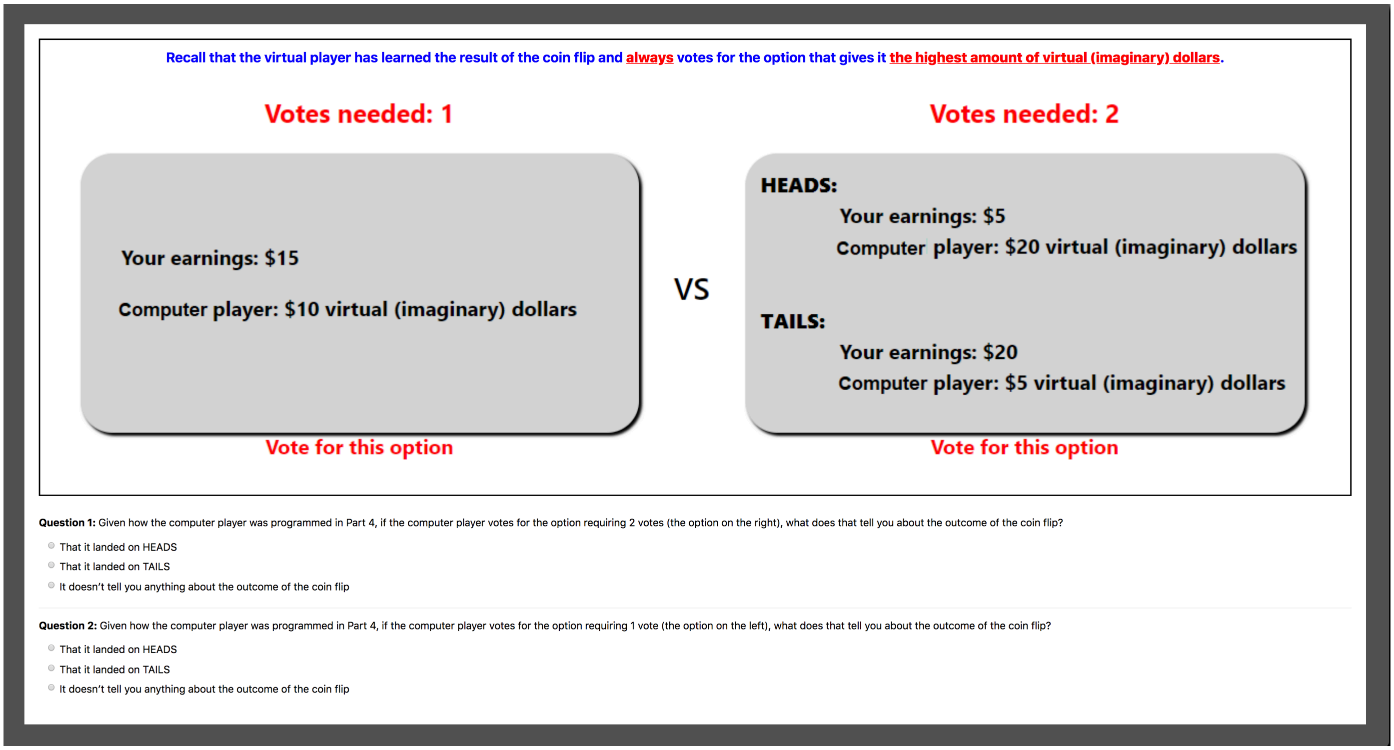


Figure 17: Part 5, Contingent-Reasoning Understanding questions (questions 1-2).

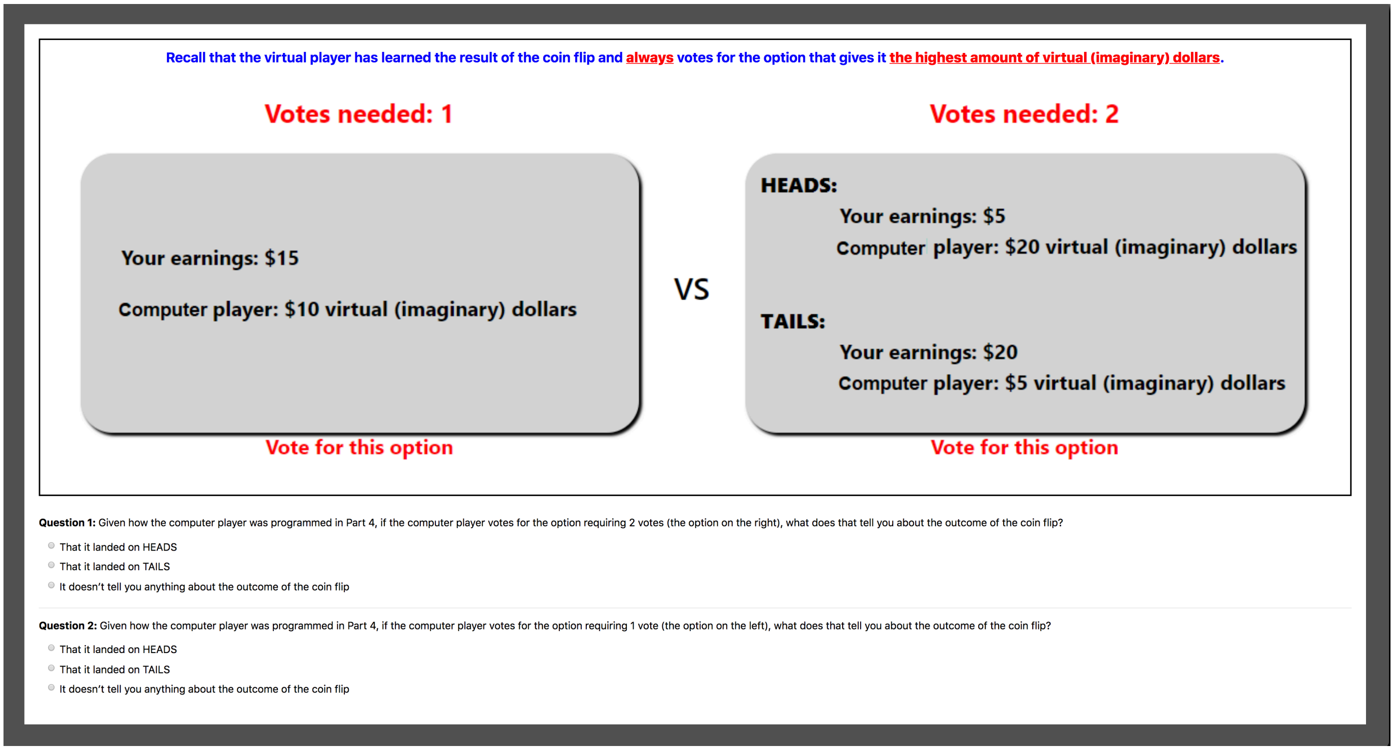


Figure 18: Part 5, Contingent-Reasoning Understanding questions (question 3).

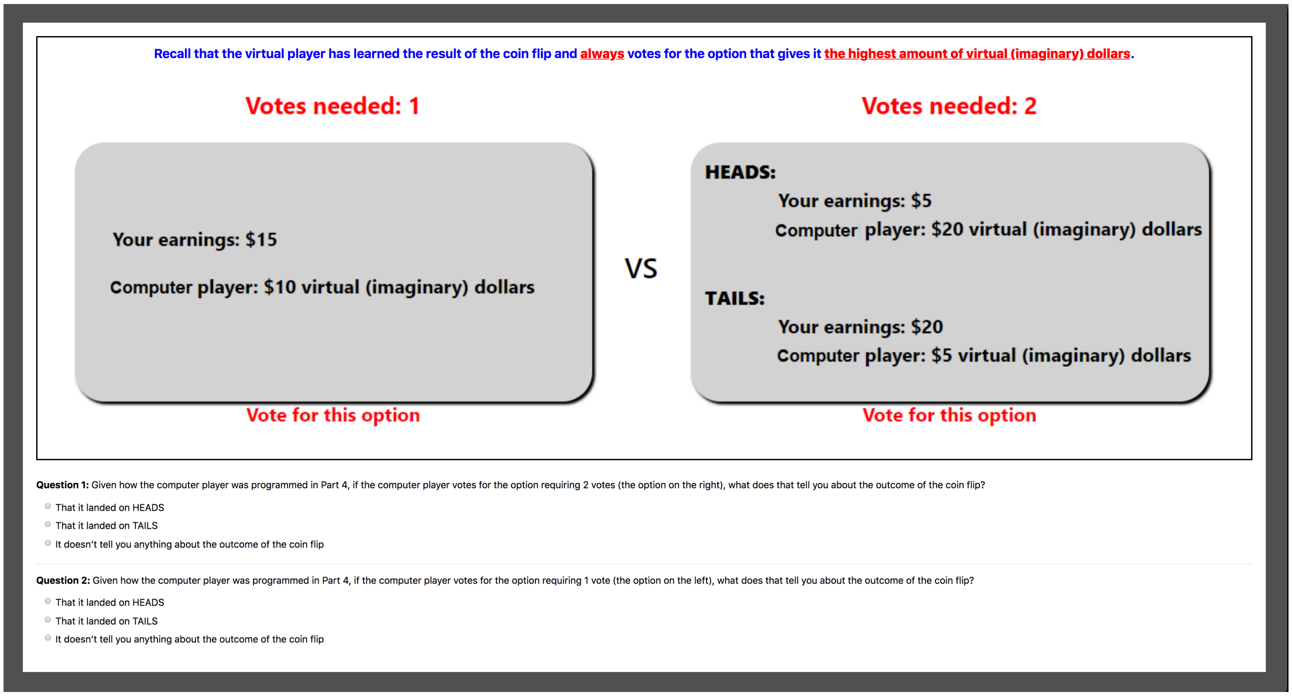


Figure 19: Part 5, Contingent-Reasoning Understanding questions (question 4).