(Re-) Inventing the Traffic Light: Designing Recommendation Devices for Play of Strategic Games*

Mikhail Anufriev[a] John Duffy[b] Valentyn Panchenko[c]
Benjamin Young[a]

September 12, 2023

[a] *Department of Economics, University of Technology Sydney*
[b] *Department of Economics, University of California, Irvine*
[c] *Economics, UNSW Business School, University of New South Wales*

Abstract

We present the results of a novel experiment investigating individuals' ability to offer incentive-compatible recommendations for strategic games. Subjects designed recommendation devices for five canonical 2×2 games played by Bayesian, expected-utility maximizing robots to achieve Pareto efficiency and fairness. Most subjects succeeded in achieving this objective in Matching Pennies and Battle of the Sexes, but only a minority found the desirable device in Prisoner's Dilemma and the two Chicken games. However, the vast majority of subjects designed an incentive-compatible recommendation device for the Chicken games. Subjects failed to recognize that strategic incentives meant the socially-efficient outcome could never be recommended in Prisoner's Dilemma. Our approach requires participants to use a holistic approach to equilibrium reasoning. Our findings suggest that equilibrium reasoning is most challenging for individuals when strategic incentives conflict with cooperative outcomes.

Keywords: Mechanism Design, Recommendations, Correlated Equilibrium, Nash Equilibrium, Experimental Economics.

JEL Codes: C72, C91, D82, D91.

*We thank Timothy Cason, Nick Feltovich, Gigi Foster, Dan Friedman, Michael McBride, Jean Paul Rabanal, Tridib Sharma, John Wooders, participants of the seminars at Chapman University, UCI, and NYU Abu-Dhabi for their comments that helped us to write this paper. The experiments were preregistered at the AEA RCT Registry (AEARCTR-0009706). The authors acknowledge the support of the Australian Research Council's Discovery Project funding scheme (project DP200101438).

1 Introduction

How good are people at equilibrium reasoning? Both the notion of Nash equilibrium (Nash Jr, 1950) and its more general counterpart, correlated equilibrium (Aumann, 1974, 1987), are important solution concepts in game theory that form the basis of a multitude of predictions and rationalizations of behavior in applied economic theory. In one interpretation of correlated equilibrium due to Myerson (1991), a mediator designs a device that makes private but correlated recommendations to each player in the game.[1] The recommendation device constitutes a correlated equilibrium if every player is always willing to follow every recommendation that is made by the device. We design a novel experiment to test the capacity of individuals to serve as the *designer* of such recommendation devices. That is, we investigate individuals' ability to coordinate the behavior of players in games towards desirable outcomes while taking into account the strategic incentives of these players to follow recommendations. Our objective is to understand whether individuals can succeed in this task across a variety of games that require different types of strategic reasoning. As such, our design sheds light on which aspects of equilibrium reasoning people understand versus those they struggle to comprehend.

Consider the following motivating example. Two cars are approaching an intersection along perpendicular roads. Conditional on the other car waiting, the driver of each car prefers to go through the intersection rather than wait. The worst outcome, however, is that both drivers go through the intersection at the same time, resulting in an accident. The issue is that neither driver has the ability to see the other and, as such, has no way to communicate who should go through the intersection while the other waits. That is, the drivers cannot coordinate their behavior. Now suppose that the city installs a traffic light that, 50% of the time displays green for one driver and, simultaneously, displays red to the other. Given a social norm that green means Go

[1] Myerson (1991), p. 250 refers to the mediator as "a person or machine that can help the players communicate and share information." Similarly, Gintis (2014), p. 142 refers to such a designer as "the choreographer," – a "new player" in the "augmented" game who issues a "directive".

while red means Stop, it follows that the drivers are willing to follow the recommendations of the traffic light *at all times* so that one driver is always passing through the intersection while the other waits. In other words, the traffic light coordinates the behavior of the drivers to achieve a fair and efficient outcome. In this example, the traffic light constitutes a recommendation device that implements a particularly desirable correlated equilibrium.

Most experimental research to date has focused on placing individuals in the role of a player in the game (i.e., a driver) and has tried to understand whether they are willing to follow or learn to follow the recommendations produced by correlating devices (i.e., the traffic light). Instead, we are the first to investigate whether subjects are able to actually design a recommendation device in order to coordinate the behavior of players within a given game (i.e., create the traffic light). As such, our research sheds light on whether individuals are able to coordinate behavior toward desirable outcomes while acknowledging the role that strategic incentives play in ensuring that recommendations are actually followed.

One may ask why we focus on this *information design problem* instead of the more traditional question of how subjects react to different information design setups. Our answer is that evidence of the ability to properly design the information structure presumes an understanding of how subjects would react to those structures and thus provides a stronger, more encompassing evaluation of the relevance of correlated and Nash equilibria. Tests of such equilibria that are conditioned on a particular signal structure are interesting but incomplete since they rely on an exogenous external event space and it is not clear how that external device may have arisen in the first place. For instance, the traffic light was designed by human engineers. Another important aim that our approach enables is a better sense of the empirical relevance of correlated versus Nash equilibria. The design of a correlation device requires mastering three interrelated concepts: strategic incentives, randomness, and correlation, and we chose our games with these three concepts in mind. Our approach allows us to investigate which of these aspects of equilibrium reasoning are most difficult to grasp.

Our experimental design is as follows. Experimental participants design recommendation devices for five games, having five attempts to design a device for each game. Each device provides probabilistic, private, potentially correlated recommendations to pairs of *robot players* playing each game. These robot players are Bayesian expected-utility maximizers who decide whether adhering to the recommendations from the participant's device aligns with their self-interest.[2] Thus, a device for which all recommendations are followed implements a correlated equilibrium.

We are interested in the subjects' ability to design a device that achieves a correlated equilibrium, and, conditional on this, results in an outcome of a game that is fair and efficient. This is accomplished as follows. First, subjects earn a positive payoff from their device if and only if all of its recommendations are followed by the robot players. Second, the payoff earned from a device for which all recommendations are followed is increasing in the minimum expected payoff across the robot players.[3] As all the games we consider are symmetric, this means that the *desirable correlated equilibrium* (DCE), which if implemented by a subject's device maximizes their payoff, is a correlated equilibrium that is both Pareto efficient and fair. We focus on this selection rule from the set of correlated equilibria as it is naturally applicable in practice. For example, the traffic light is a recommendation device that achieves this social objective in the context of safely managing traffic flow. This is because it is Pareto efficient (cars from only one direction are traversing the intersection at a time), and it is fair (cars from different directions are afforded the opportunity to cross the intersection).

There are five games for which participants are asked to design recommendation devices: Prisoner's Dilemma, Matching Pennies, Battle of the Sexes, and two

[2] By having participants provide recommendations to rational robots following prespecified rules, we abstract from the issues inherent in studies that test whether subjects are willing to follow correlated-equilibrium recommendations. These issues include strategic uncertainty regarding others' actions and social preferences. Instead, we test the extent to which subjects can engage in the equilibrium reasoning needed for adequate coordination in games.

[3] In this sense, subjects are provided the *Rawlsian criterion* (see Rawls, 1971) – to make the worst-off player as well off as possible, – as a secondary objective in selecting a correlated equilibrium. Subjects attempt to satisfy this criterion subject to the constraint that the device is incentive compatible (i.e., constitutes a correlated equilibrium).

versions of Chicken that differ in both the set of correlated equilibria and their desirable correlated equilibrium. These games are chosen because the structure of their DCE differs in important aspects of equilibrium reasoning, such as (1) the necessity of randomization, (2) the degree of correlation required, and (3) the alignment of strategic incentives with the social objective of fairness and efficiency. For example, the DCE in Prisoner's Dilemma is the unique Nash equilibrium in pure strategies, which does not require randomization or correlation, and has payoff incentives that are at odds with efficient outcomes. In contrast, the DCE in Battle of the Sexes involves perfectly-correlated randomization between its two pure-strategy Nash equilibria and has payoff incentives that are perfectly aligned with the objective of fairness and efficiency. These differences allow us to shed light on the aspects of equilibrium reasoning that are most difficult for individuals.

We find significant heterogeneity across games in terms of the ability of subjects to design recommendation devices that implement the desirable correlated equilibrium. The majority of subjects are able to design the DCE in Battle of the Sexes and Matching Pennies (79% and 73% of subjects, respectively) within the five attempts provided, while only a minority of subjects do so in Prisoner's Dilemma and the two Chicken games (37%, 18% and 30% of subjects, respectively). In both Chicken games, however, subjects generally find correlated equilibria that partially meet the experiment's objective. That is, they design devices for which all recommendations are followed even if their correlated equilibria are not fair and Pareto efficient. Using a measure based on the distance between different distributions, we see that subjects design devices that are close to what is desirable in all games except Prisoner's Dilemma. Thus, our findings suggest that Prisoner's Dilemma is the most difficult game for subjects to design incentive-compatible recommendation devices.

We also explore in detail the devices that subjects design at the individual level. We find that subjects initially design devices that put a lot of weight on action profiles that are fair and efficient for each game. In Matching Pennies and Battle of the Sexes, this constitutes a successful strategy. However, in Prisoner's Dilemma and both versions of Chicken, doing so does not respect the strategic incentives of the robot players. By their final attempt, however, the modal device corresponds to the DCE in each of the five games. However, many subjects are still putting too much weight on the fair and efficient action profile in Prisoner's Dilemma, while they are, overall, underestimating the extent to which they can recommend such profiles in Chicken. Thus, it seems that they understand better, at least qualitatively, the role of strategic incentives in Chicken relative to Prisoner's Dilemma.

Our results suggest that subjects in our experiment are comfortable with randomizing and correlating their recommendations across games. Instead, subjects appear to struggle most when strategic incentives are not aligned with socially desirable outcomes. By this, we mean situations in which the set of correlated equilibria does not include the fair and Pareto efficient distribution of player actions when incentive constraints are not imposed. Indeed, subjects exhibit maximal success in games for which the socially desirable outcome is incentive compatible (Matching Pennies and Battle of the Sexes), moderate success when strategic incentives are only weakly associated with the socially desirable outcome (both versions of Chicken), and low success when strategic incentives are unaligned with the socially desirable outcome (Prisoner's Dilemma). This suggests that individuals do not separate their cooperative mindset from the non-cooperative mindset of the robot-players who respond to these recommendations. This is particularly interesting since one cannot conclude that this is due to the usual factors used to explain cooperative behavior in Prisoner's Dilemma, such as social preferences. Moreover, it is striking since the strategic incentives pushing behavior away from the socially desirable outcome in the Prisoner's Dilemma are very strong: players have a strictly dominant strategy in this setting.

2 Related Literature

There are a number of papers that experimentally investigate correlated behavior in games. These papers find that individuals tend to correlate their behavior when preplay communication is permitted (Moreno and Wooders, 1998), follow correlated-equilibrium recommendations when those improve upon Nash equilibrium outcomes (Duffy and Feltovich, 2010), and may learn to correlate their behavior in the absence of coordinating devices (Friedman et al., 2022). The literature also finds, however, that the propensity for subjects to follow recommendations that constitute a correlated equilibrium is mediated by factors such as strategic uncertainty (Cason and Sharma, 2007), the directness of recommendations (Duffy et al., 2017), payoff asymmetries (Anbarcı et al., 2018), whether recommendations are private or public (Bone et al., 2013), and whether subjects must commit to the coordinating device as a whole (Georgalos et al., 2020). These papers focus mostly on the willingness of individuals to adhere to the recommendations made by correlation devices. By contrast in this paper, we ask whether individuals are able to design the correlation device themselves.

An exception to studies that investigate subjects' willingness to follow correlated-equilibrium recommendations is the work of Cason et al. (2022). In their paper, experimental subjects are asked to predict the distribution over actions that will be generated by other individuals playing that game. They find that subjects tend to believe that play will be correlated even though the people playing the games do not have access to any correlating device. In contrast, instead of asking subjects to predict what *will* happen in the game, we ask subjects to design a coordination device that implements what they believe *should* happen in the game. Moreover, our subjects design devices for rational robot players to avoid the strategic uncertainties created by human players.

There is also a related literature on mechanism design experiments, surveyed in Chen and Ledyard (2010). These experiments generally involve evaluation of subjects' play under various imposed mechanisms with most applications concerning public good provision, auctions, contract theory, matching markets, and prediction markets, among others. In this literature, it is typically *not* the case that subjects are placed in the designer role as in our study. Still, we think that for simple games, such as those that we study here, this is not an unreasonable approach. This is because it reveals the type of designs that non-expert designers find most reasonable to implement.

Our work is also related to the literature on experiments regarding information design and Bayesian persuasion (Kamenica and Gentzkow, 2011). These papers ask whether subjects can use information structures that successfully correlate behavior with some payoff-relevant state of the world (Aristidou et al., 2019; Au and Li, 2018; Fréchette et al., 2022; Wu and Ye, 2021; Ziegler, 2022). In our paper, we focus on complete information games (i.e., a single payoff-relevant state) and subjects seek to create signal structures or recommendations that correlate the behavior of players in that complete information game (i.e., implement correlated equilibria).

Finally, our work can be related to that which explores the ability of individuals to draw inferences from correlated signals and the notion of correlation neglect (Enke and Zimmermann, 2019; Hossain and Okui, 2021). These issues are more likely to affect the players in the game who must decide whether to follow correlated recommendations made privately. We, instead, ask the reverse question: can subjects actually create the correlated signals that inferences must be drawn from? As such, our subjects need to think about signal correlation *hypothetically* rather than draw inference from *actual* signal realizations.[4]

3 Theoretical Framework

A finite game of complete information can be described by the tuple

$$G = (N, (A_i)_{i \in N}, (u_i)_{i \in N}),$$

where N is a set of players, A_i is a finite action set for player $i \in N$, and $u_i : A \to \mathbb{R}$ is the utility of player i defined over the set of action profiles $A \equiv \prod_{i \in N} A_i$. Let $\Delta(A)$ be the set of probability distributions over A. For a game G, we can define a distribution over action profiles $\mu \in \Delta(A)$ to be a correlated equilibrium as follows.

[4]This is an important distinction as there is evidence that thinking on hypothetical events is fundamentally different from extracting information from actual events (Esponda and Vespa, 2014).

Definition 1. *Distribution $\mu \in \Delta(A)$ is a **correlated equilibrium** for game G if*

$$\sum_{a_{-i}} \mu(a_i, a_{-i})[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})] \geq 0$$

for all $i \in N$, $a_i, a_i' \in A_i$.

In order to interpret a correlated equilibrium, one can think of μ as a statistical summary of some recommendation device, where action profile a is drawn with probability $\mu(a)$ and each player i is provided with only their own recommended action a_i from profile a. Then, μ constitutes a correlated equilibrium if and only if each player i finds it optimal to follow every action recommendation they receive given that all other players also follow their own recommendations from μ.[5] Note that a Nash equilibrium is a special case of a correlated equilibrium in which μ is the product of its marginal distributions across players; that is, action recommendations are made *independently* to the players in the game.

In our experiment, subjects will design recommendation devices in an attempt to implement correlated equilibria. Since the sizes of sets of correlated equilibria vary across games, we focus on one particular correlated equilibrium in each game. Specifically, we are interested in the solution of the following optimization problem:

$$\begin{gathered} \max_{\mu \in \Delta(A)} \min_{i \in N} \sum\nolimits_{a \in A} \mu(a) u_i(a)\,, \\ \text{s.t. } \mu \text{ is a correlated equilibrium.} \end{gathered} \tag{1}$$

That is, we are interested in the correlated equilibrium that satisfies the Rawlsian criterion of maximizing the minimum average payoff across players (Rawls, 1971). Each of the games that our subjects design recommendation devices for is symmetric so that the problem in (1) is equivalent to finding the correlated equilibrium that is both efficient (i.e., surplus maximizing) and fair (i.e., equal surplus shares

[5]There are alternative theories of why observed behavior in games might constitute a correlated equilibrium that do not require the interpretation of a recommendation device. For example, correlated equilibria can simply arise as Bayes' Nash equilibria when players can condition on extraneous signals, or it can result from simple adaptive procedures with repeated play (Hart and Mas-Colell, 2000). We focus on the interpretation of a recommendation device as this is a natural role to place an individual in to investigate their ability to implement correlated equilibria.

are allocated to each player). We call this distribution the **desirable correlated equilibrium** (DCE) and denote it by μ^*.

We focus on the correlated equilibrium that is fair and Pareto efficient as it is a natural objective in practice. For example, the traffic light is an example of a recommendation device that implements a fair and efficient outcome, as discussed in the Introduction. Moreover, there is experimental evidence that individuals are more likely to follow correlated-equilibrium recommendations when there are efficiency gains relative to Nash equilibrium (Duffy and Feltovich, 2010) and asymmetries in payoffs are minimized Anbarcı et al. (2018) (i.e., they are fair). Consequently, recommendation devices should be more effective at coordinating real-world behavior if they satisfy the criterion in (1).

Throughout the paper, we will also often refer to the **socially desirable outcome** which is the unconstrained solution to problem (1). In other words, this is the probability distribution $\mu \in \Delta(A)$ that is fair and efficient but not necessarily a correlated equilibrium. We denote this distribution by μ^{SD}. The socially desirable outcome does not coincide with the DCE, whenever strategic incentives force at least one player not to follow at least one recommendation. In such situations, we will say that the strategic incentives of the game are not perfectly aligned with the socially desirable outcome.

4 Experimental Design

Our experiment employed a within-subjects design and was coded using oTree (Chen et al., 2016). Each subject designed recommendation devices for five different games, presented in random order. For each game, subjects were given five consecutive attempts (which we call rounds) to design their device.

We recruited 100 undergraduate students from the University of California, Irvine, pursuing various programs of study. Thus, we collected, in total, 2500 designed recommendation devices (100 participants $\times$ 5 games $\times$ 5 rounds in each game).

The study was gender balanced with 54 females and 46 males. Subjects had no prior experience with our experiment; each subject participated in a single session with five games. Payments were determined by converting points earned by their recommendation devices into USD (more detail provided in Section 4.3). In addition, subjects earned a 10 USD show-up payment. Average total subject earnings were 27.64 USD for an experiment lasting, on average, 90 minutes (including approximately 30 minutes spent reading instructions and 15 minutes spent for the quiz).

At the end of the experiment, we collected additional information from the subjects in order to investigate whether there were any notable correlates with performance in the design task. This additional information included standard demographic information, field and year of study, performance on a five-question version of a cognitive reflection test (CRT) (Frederick, 2005), a measure of strategic reasoning related to the hypothetical play of the 11-20 game (Arad and Rubinstein, 2012), and the self-reported strategy utilized by subjects.[6]

4.1 The Design Problem

Subjects designed recommendation devices for a selection of 2×2 games which take the general normal form in Fig. 1. Each player in the game has two actions, either Red (denoted R) or Blue (denoted B). We chose colors as this provided a nice visual representation of the recommendation device that subjects must design.

P1/P2	Red (R)	Blue (B)
Red (R)	(u_1^{RR}, u_2^{RR})	(u_1^{RB}, u_2^{RB})
Blue (B)	(u_1^{BR}, u_2^{BR})	(u_1^{BB}, u_2^{BB})

Figure 1: Normal game form for an arbitrary 2×2 game where both Player 1 and Player 2 choose between the actions Red and Blue.

An arbitrary recommendation device for a 2×2 game can be summarized as a table presented in Fig. 2. One can interpret $\mu_{a_1 a_2} \equiv \mu(a_1, a_2)$ in this table to be the

[6] For the precise wording of the CRT questions and the 11-20 question, see Online Appendix B.

P1/P2	Red (R)	Blue (B)
Red (R)	μ_{RR}	μ_{RB}
Blue (B)	μ_{BR}	μ_{BB}

Figure 2: Table representation of an arbitrary recommendation device for a 2×2 game in which both players choose between actions Red and Blue.

probability that the device jointly recommends actions $a_1 \in \{R, B\}$ and $a_2 \in \{R, B\}$ to player 1 and player 2 respectively. The process that subjects used to design this device is described in detail in the next section.

Subjects were told that their goal is to make recommendations to "intelligent" robot players that would follow recommendations *only if* it is in their best interest to do so (i.e., they have a strict preference to follow recommendations or are indifferent).[7]

4.2 Designing the Device

In this section, we describe in detail the protocol that subjects followed in order to design their recommendation devices for each game. Subjects were provided on the screen with a bucket of 24 balls, where each ball was labeled with '1' on one half and '2' on the other half.

In order to design the device, subjects were required to color in each half of these balls with either the color red or blue. They were told that for each pair of robot-players, a random ball is selected from the bucket and is split in half, with the half labeled '1' given to robot-player 1 and the half labeled '2' given to robot-player 2. The idea, which was explained to them, is that the color of the half given to a robot-player would be the action recommended to that robot-player.

The balls were colored in by the subjects using sliders. First, subjects used a slider to determine, for the half labeled '1', how many to color red and how many to color blue for all 24 balls. After doing this, subjects then used two additional

[7] Essentially we explained in layman's terms that the robot players are Bayesian expected-utility maximizers. See page 2 of the experimental instructions in the Online Appendix A for the precise explanation. Pages 6-8 of the instructions provide examples of the robots' reasoning.

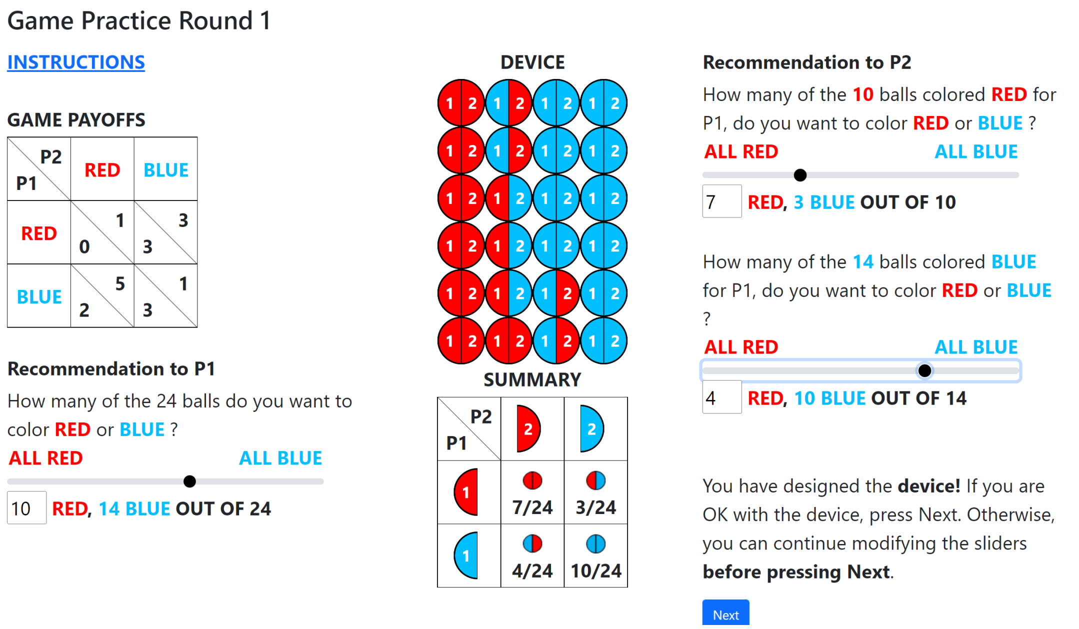


Figure 3: Example of the decision screen used by subjects to input a recommendation device. The game shown here was used in the instructions and for practice rounds.

sliders to color the half labeled '2': one for balls that have half '1' colored red and one for balls that have half '1' colored blue. After all balls are colored in, the design of the device is completed and subjects are prompted as to whether they would like to submit their design for evaluation. Before confirming the submission of their device, they were able to continue to make any alterations they saw fit. See Fig. 3 for an example of the decision screen with a designed device. Note that the game that the subjects were designing the recommendation device for, as well as a summary of the device in the form presented in Fig. 2 were also provided to subjects.

While limited by the fact that the number of balls to color is finite, this device still allowed subjects to implement a large number of probability distributions (there were, precisely, 2925 probability distributions that subjects could feasibly design). Hence, it was unlikely that subjects found the desirable distribution of colored balls at random in the five attempts afforded for each game.

4.3 Incentivization

For each game, subjects were incentivized to design a recommendation device that solves problem (1), i.e., a device that implements the correlated equilibrium in which the minimum payoff across the robot players is as large as possible. We did this by ensuring that the points subjects earned from designing a particular device were incentive-compatible for this objective.

Specifically, it was explained to subjects that a large number of pairs of robot-players 1 and 2 would be drawn to play this game against each other.[8] They were told that if any single robot-player from all of these draws did not want to follow one of the recommendations made by their device, then they would receive no points for this device. Moreover, for each designed device in which *all* recommendations were followed (i.e., constituted a correlated equilibrium), they were told that they would earn a number of points equal to the minimum of the average payoff across all robot-players 1 and robot-players 2. It is clear that, together, these two conditions for payment are consistent with incentivizing subjects to solve problem (1), that is to submit the DCE μ^* for the game. Note that subjects receive a *positive* payoff for a device if and only if all its recommendations are followed. In this sense, subjects are also incentivized to target the set of correlated equilibria as a whole.

At the end of the experiment, a random game was chosen and subjects were paid for the device that earned the *highest* number of points for that game taken across all 5 rounds. We did this to ensure that subjects had sufficient incentives to continue experimenting in order to find the device that implemented the desirable correlated equilibrium.[9] Subjects were paid at an exchange rate of 6 USD per point earned in addition to their show-up fee of 10 USD.

[8] Even though we explained how the device would be utilized in this dynamic way, we emphasized that robots would treat each round as a one-shot game when deciding whether to follow recommendations or not. See page 2 of the experimental instructions in the Online Appendix A for further details on how this was explained to subjects.

[9] If we had chosen a random round within a random game, then subjects may have decided to stop experimenting once they found *some* correlated equilibrium (not necessarily the DCE) in order to maximize the probability of receiving some payment at the end of the experiment.

4.4 Feedback

We provided feedback to subjects between rounds. In particular, for a given inputted device, we told them whether or not all recommendations from their device were followed (i.e., whether a correlated equilibrium was reached or not). If a correlated equilibrium was reached, we additionally told them the number of points their device earned and whether it could be improved upon to earn more points (see Figs. B1 and B2 in Online Appendix C). This, in conjunction with the fact that only the best device across the five rounds in a given game was considered for payment, incentivized subjects to experiment with their devices.

If a correlated equilibrium was not reached, we provided subjects with a ball for which one of the robot players did not follow the recommendation, chosen in proportion to their designed device. For this ball, we explained that one of the players did not want to follow their recommendation and some reasoning as to why this was the case, including reinforcement of the idea that subjects must think about their device as a whole to ensure incentive compatibility (see Fig. B3 in Online Appendix C). We chose to provide only one ball at a time as feedback in order not to overload subjects with information.

4.5 Reinforcement of the Experimental Task

We went to great lengths to ensure that subjects understood the experimental task. In particular, we wanted to make sure that they understood how to use the interface we created to design recommendation devices, how robot-players would respond to their devices, and how their payment would be determined.

To this end, we provided detailed written instructions that were read aloud at the start of the experiment (see Online Appendix A). These instructions outlined details on all important components of the experiment, provided examples of recommendation devices for a specific game outside of those experimentally tested, and also described an algorithm that subjects could use to implement any particular recommendation device they wanted in the form provided in Fig. 2. Subjects were then required to complete a nine-question quiz which involved answering questions about the salient aspects of the experiment and also involved designing devices explicitly. Importantly, subjects were not allowed to begin the actual experiment until this quiz was completed correctly and we provided detailed feedback for each answer. Finally, subjects were given three practice rounds with the same game that we utilized as an example in the instructions. All these opportunities, combined, gave subjects a degree of comfort with the experimental task that should have helped to ameliorate noise coming from subjects simply not being able to use the interface we provided.

4.6 Games Considered in the Experiment

Subjects designed recommendation devices for five games: Prisoner's Dilemma, Matching Pennies, Battle of the Sexes, and two versions of the Chicken game. The exact formulation of each game is presented in the left column of Table 1. In addition to the matrix of each game, we report the volume of the set of all correlated equilibria (as a fraction of the volume of the set of probability distributions over action profiles), and the payoff in the desirable correlated equilibrium, i.e., in the probability distribution that solves problem (1). We list here the DCE for each game as we will benchmark our results to these distributions.[10]

(i) **Prisoner's Dilemma:** The DCE is $\mu^*_{RR} = 1$ and $\mu^*_{RB} = \mu^*_{BR} = \mu^*_{BB} = 0$.

(ii) **Matching Pennies:** The DCE is $\mu^*_{RR} = \mu^*_{RB} = \mu^*_{BR} = \mu^*_{BB} = 1/4$.

(iii) **Battle of the Sexes:** The DCE is $\mu^*_{RB} = \mu^*_{BR} = 1/2$, and $\mu^*_{RR} = \mu^*_{BB} = 0$.

(iv) **Chicken, version 1:** The DCE is $\mu^*_{RB} = \mu^*_{BR} = \mu^*_{BB} = 1/3$, and $\mu^*_{RR} = 0$.

(v) **Chicken, version 2:** The DCE is $\mu^*_{BB} = 1/2$, $\mu^*_{RB} = \mu^*_{BR} = 1/4$, and $\mu^*_{RR} = 0$.

[10] We derive these equilibria in Online Appendix D.

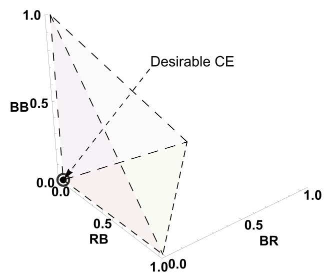
<table>
<tr><th>Game</th><th>Distribution Space</th><th>Payoff Space</th></tr>
<tr><td>

Prisoner's Dilemma

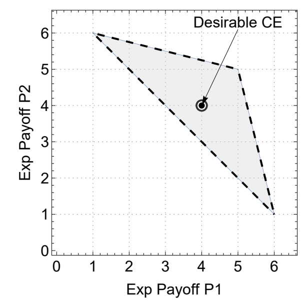
<table><tr><th>PD</th><th>Red</th><th>Blue</th></tr><tr><td>Red</td><td>(4, 4)</td><td>(6, 1)</td></tr><tr><td>Blue</td><td>(1, 6)</td><td>(5, 5)</td></tr></table>

CE set volume: 0
Desirable CE payoff: 4

</td><td>



</td><td>



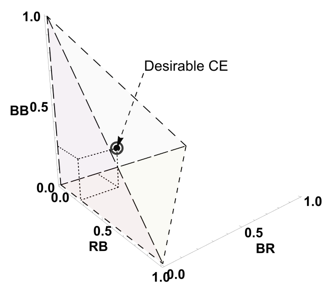
</td></tr>
<tr><td>

Matching Pennies

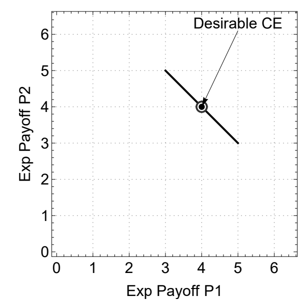
<table><tr><th>MP</th><th>Red</th><th>Blue</th></tr><tr><td>Red</td><td>(5, 3)</td><td>(3, 5)</td></tr><tr><td>Blue</td><td>(3, 5)</td><td>(5, 3)</td></tr></table>

CE set volume: 0
Desirable CE payoff: 4

</td><td>



</td><td>



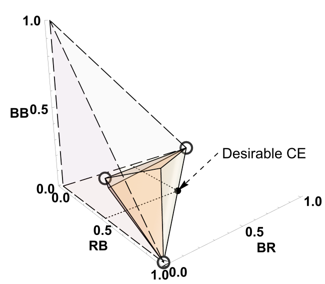
</td></tr>
<tr><td>

Battle of the Sexes

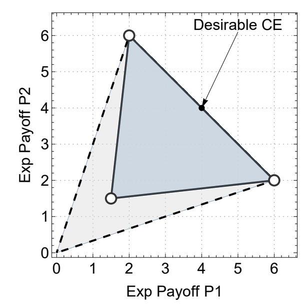
<table><tr><th>BoS</th><th>Red</th><th>Blue</th></tr><tr><td>Red</td><td>(0, 0)</td><td>(6, 2)</td></tr><tr><td>Blue</td><td>(2, 6)</td><td>(0, 0)</td></tr></table>

CE set volume: 0.118
Desirable CE payoff: 4

</td><td>



</td><td>



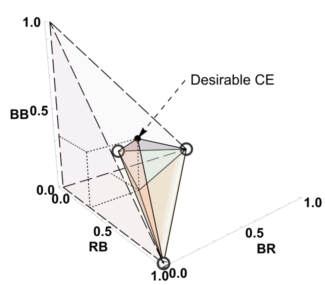
</td></tr>
<tr><td>

Chicken - ver 1

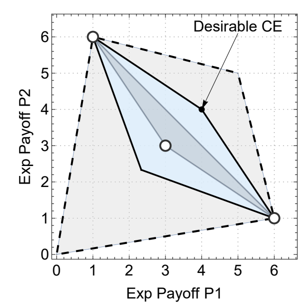
<table><tr><th>C1</th><th>Red</th><th>Blue</th></tr><tr><td>Red</td><td>(0, 0)</td><td>(6, 1)</td></tr><tr><td>Blue</td><td>(1, 6)</td><td>(5, 5)</td></tr></table>

CE set volume: 0.167
Desirable CE payoff: 4

</td><td>



</td><td>



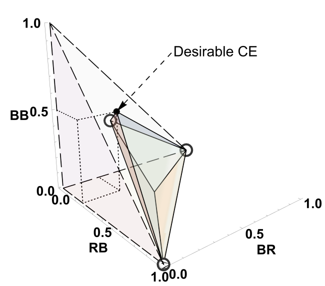
</td></tr>
<tr><td>

Chicken - ver 2

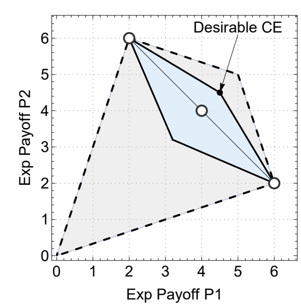
<table><tr><th>C2</th><th>Red</th><th>Blue</th></tr><tr><td>Red</td><td>(0, 0)</td><td>(6, 2)</td></tr><tr><td>Blue</td><td>(2, 6)</td><td>(5, 5)</td></tr></table>

CE set volume: 0.144
Desirable CE payoff: 4.5

</td><td>



</td><td>



</td></tr>
</table>

Table 1: Games used in the experiment.

Table 1 also provides a graphical representation of each game in both the joint distribution space (middle panels) and the payoff space (right panels). In these plots, white discs correspond to the Nash equilibria of the game, and the black dot shows the DCE. The distribution space is a three-dimensional simplex that we represent in coordinates $(\mu_{RB}, \mu_{BR}, \mu_{BB})$. Therefore, the origin corresponds to the full weight given to (R, R) recommendation. This is the Nash equilibrium of the Prisoner's Dilemma, for example. In the first two games, there is a unique correlated equilibrium, and so their sets are singletons with zero volume. For the last three games, a set of correlated equilibria is a proper polytope; its edges are drawn in black and the faces are colored. We also indicated projections of the DCE to the axis by dashed lines. The space of expected payoffs of two robot players is two-dimensional and thus provides a better readable illustration. However, it should be emphasized that these are projections from the distribution space. In all plots, we can see the set of feasible payoffs in the game (the light-grey shaded area) and the set of correlated equilibria in each game (the light-blue shaded area with a thick solid boundary). We also show the convex hull of Nash equilibrium payoffs (the dark-blue shaded area). In both Chicken games, the convex hull is a proper subset of the set of correlated equilibria, but in the Battle of the Sexes, these two sets coincide.

Online Appendix D shows how to design the device for each game using our interface in order to earn a maximal payoff. Notice that, for each game, the DCE corresponds to the correlated equilibrium that is both Pareto efficient and fair in the sense that it is on the upper frontier of correlated equilibrium payoffs and the expected payoffs of both robot-players are the same.

The DCE varies across games in a number of aspects: (a) whether randomization is required, (b) the degree of correlation required, and (c) the extent to which strategic incentives are aligned with the socially desired outcome, μ^{SD}, defined as the solution of the *unconstrained* problem (1). These differences are described in Table 2.

All games require randomization except for Prisoner's Dilemma in which the unique correlated equilibrium is the unique Nash equilibrium in pure strategies. Nei-

	Randomization	Correlation	Alignment of Incentives
PD	No	None	Strongly Unaligned
MP	Yes	None	Aligned
BoS	Yes	Perfect	Aligned
C1	Yes	Partial	Weakly Aligned
C2	Yes	Partial	Weakly Aligned

Table 2: Differences in the types of reasoning required to solve for the DCE across games.

ther Prisoner's Dilemma nor Matching Pennies require correlation since the DCE in these games is also the Nash equilibrium. Instead, Battle of the Sexes requires perfect correlation (i.e., can be implemented using a public randomization device) while the two Chicken games require more nuanced, partial correlation (i.e., private recommendations are required). Finally, strategic incentives are aligned with the socially desirable outcome in Matching Pennies and Battle of the Sexes in the sense that their DCEs are also socially desirable (i.e., the constrained and unconstrained version of (1) have the same solution). Instead, in Prisoner's Dilemma and both versions of Chicken the socially desirable outcome is $\mu^{\text{SD}}_{BB} = 1$, which is *not* a correlated equilibrium in any of these games. However, we say that strategic incentives are *weakly aligned* with the socially desirable outcome in Chicken as the DCE puts positive weight on this profile in each of these games ($\mu^{\text{SD}}_{BB} = 1/3$ in C1 and $\mu^{\text{SD}}_{BB} = 1/2$ in C2). Instead, we say that incentives are strongly unaligned in the case of Prisoner's Dilemma since $\mu^{\text{SD}}_{BB} = 0$ in this game.[11]

5 Results

Our experimental design allows us to address three main research questions:

(1) Can individuals design incentive-compatible recommendation devices that are fair and efficient and how does the ability to do so vary across the 5 games?

[11] We provide a more quantitative measure of the extent to which strategic incentives and the socially desirable outcome are aligned in Section 5.3.3.

(2) What sorts of recommendation devices do they design for each game and how does this evolve with experience?

(3) How do the answers to (1) and (2) shed light on which aspects of strategic reasoning and equilibrium thinking are most difficult for people to understand?

We now present evidence that seeks to address these research questions in turn.

5.1 How Well Did Subjects Perform?

In this section we discuss the overall performance of subjects in the experimental task, focusing on comparing performance across games.

5.1.1 Performance Measures

We first describe three measures that we use to analyze how successful subjects were in designing their recommendation devices. We define each measure for a given game and a given round. The first two measures are based on the number of points that subjects earned with their recommendation devices, capturing the extent to which participants achieved the goals they were incentivized to. The third measure is based on the distance between distributions and complements the first two by showing how close the device that participants designed was to the DCE.

1. **Score** $\in [0, 1]$ is the number of points a participant earned, relative to the maximum number of points that could have been earned for the game. Thus, **Score** $= 1$ if the DCE of the game is implemented and **Score** > 0 as soon as a correlated equilibrium is reached.[12]

2. **TopScore** $\in \{0, 1\}$ is an indicator variable that is equal to 1 when a subject designed the DCE and is equal to 0 otherwise. Note that **TopScore** $= 1$ if and only if **Score** $= 1$, as subjects get their maximum number of points in the DCE.

[12] Normalization is useful as the maximum number of points that could be earned varied across the different games. Specifically, the second version of Chicken has a maximum payoff in points of 4.5 while all other games have a maximum payoff in points of 4.

3. **N-RMSE** $\in [0,1]$ is a normalized measure based on the Root Mean Square Error (RMSE) capturing distance between two distributions.[13] The RMSE between two distributions μ and ν is defined as

$$\text{RMSE}(\mu,\nu) \equiv \sqrt{\sum_{a\in\{R,B\}^2} (\mu(a)-\nu(a))^2}\,.$$

We will always use the following transformation of the RMSE to ensure that this performance measure is within $[0,1]$ interval and increases as μ moves closer to ν:

$$\text{N-RMSE}(\mu;\nu) \equiv 1 - \frac{\text{RMSE}(\mu,\nu)}{\max_{\mu'} \text{RMSE}(\mu',\nu)}.$$

We use N-RMSE$(\mu;\mu^*)$ to measure how close a subject's designed distribution, μ, is to the DCE, μ^*. Throughout the paper we use N-RMSE as a shorthand for this measure. Note that N-RMSE $= 1$ if and only if $\mu = \mu^*$ and N-RMSE $= 0$ if μ is maximally distant from μ^* in terms of Root Mean Square Error.

5.1.2 Aggregate Performance

We first analyze how subjects performed in the aggregate in terms of their *best attempt* at finding the DCE in each game. To this end, we fix a participant and, for each game, we compute all three performance measures for each of the five rounds. Then, for each performance measure, we take the maximum across the five rounds and average this across the sample of participants. Fig. 4 provides a graphical comparison of aggregate performance across games for each of the measures.

Inspecting Fig. 4, we see that performance is highest in Battle of the Sexes across all measures. Looking at **TopScore**, approximately 80% of subjects achieve the DCE in this game, which requires perfectly-correlated mixing between the two pure-strategy Nash equilibria of the game. Next is Matching Pennies, in which nearly 75% of participants found the DCE which coincides with full randomization over action profiles. Aggregate performance in these two games is nearly indistinguishable: for

[13] Normalization is made for the ease of comparison with the other two measures.

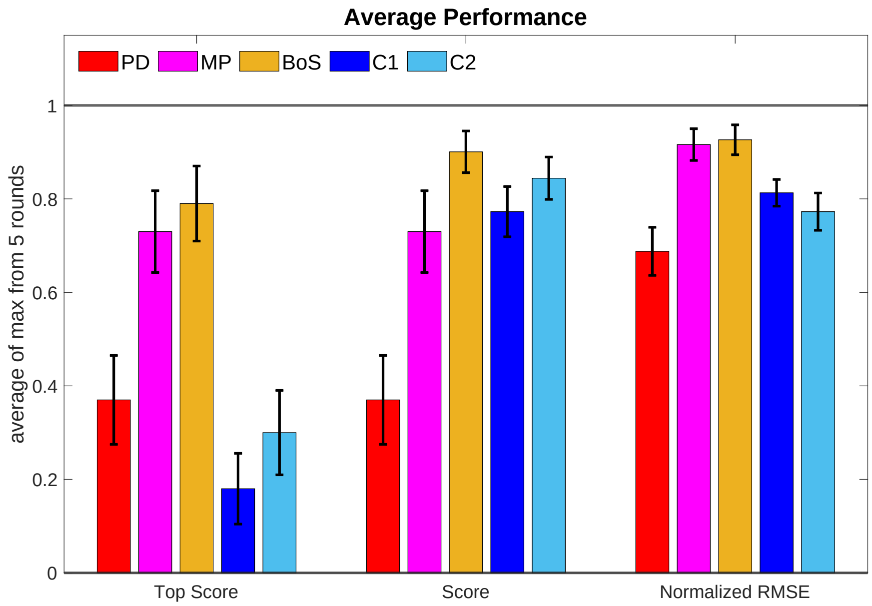


Figure 4: The average (across all subjects) of the highest in five rounds of different performance measures for each game. Error bars represent 95% confidence intervals.

two performance measures (except for **Score**), Matching Pennies is not statistically different from Battle of the Sexes.[14]

In terms of these performance measures, subjects found Prisoner's Dilemma and the two Chicken games most difficult. Looking at **TopScore**, we see that approximately 40% of subjects found the optimum in Prisoner's Dilemma, 30% of subjects found the optimum in Chicken 2, and nearly 20% of subjects found the optimum in Chicken 1. However, in Chicken games, subjects appear to have struggled more with the quantitative issue of calibrating their device precisely to meet the design objective, rather than the qualitative issue of coordinating behavior within the game. Indeed, **Score** is high in these games, implying that subjects were able to consistently find a correlated equilibrium that was relatively suitable in terms of the objective function in (1).

Our distance-based performance measure suggests that Prisoner's Dilemma was

[14]This is based on a standard two-sided t-test for the difference of two averages at the 5% significance level, which we use hereafter unless specified otherwise.

the game in which subjects exhibited the least understanding of what was required to solve the problem in a qualitative sense. This is evidenced by the fact that subjects were the most distant from the DCE in terms of their designed distributions for this game. Specifically, the average of N-RMSE is the smallest for Prisoner's Dilemma and is significantly different from that of each other game. Instead, for Matching Pennies and Battle of the Sexes, is close to one, suggesting that even when subjects could not find the DCE in these games, they were close to it. N-RMSE in the Chicken games is significantly less than in each of these games. Thus, for the distance-based measure, performance was intermediate in the Chicken games (bounded above by Battle of the Sexes and Matching Pennies but significantly better than in Prisoner's Dilemma).

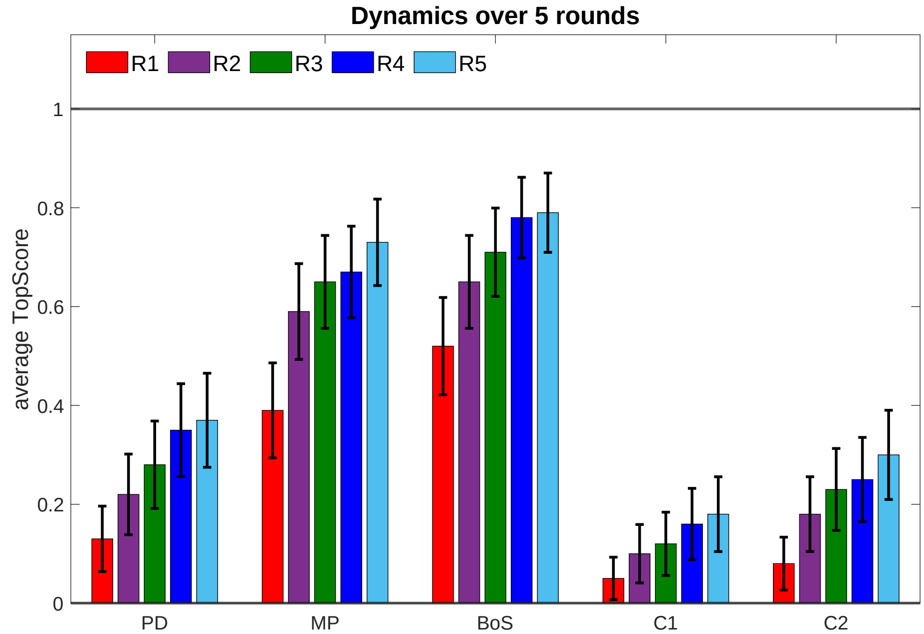
5.1.3 Performance Dynamics

We now proceed to describe how our aggregate performance measures evolved across the five rounds within each game. In order to do this, it makes sense for us to amend our raw dataset and impute the DCE device for all rounds in a given game, after a subject achieved the DCE.[15]

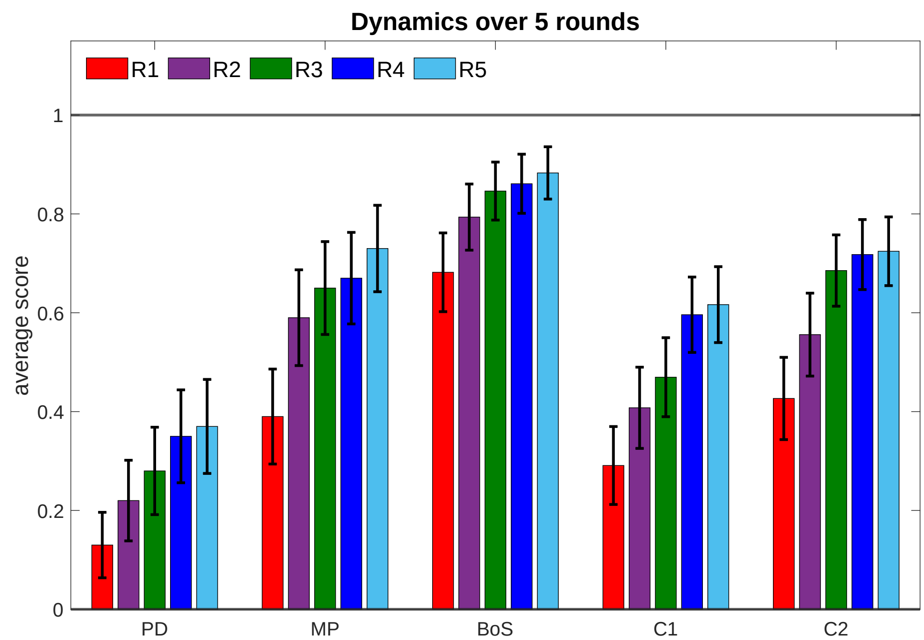
Fig. 5 displays the dynamics across rounds for **TopScore** and **Score**. Here, we see that average performance improves across rounds. Indeed, performance under both measures is significantly better in round 5 (R5) than it is in round 1 (R1) for all games. However, the improvement does appear to slow down already with only five rounds of experience: the difference between either **TopScore** or **Score** in round 5 (R5) and round 3 (R3) is not significant for any game.

Given our amended dataset, we can interpret **TopScore** in a given round to be

[15]Subjects had to complete all five rounds for a game even if they achieved the DCE in an earlier round. Whereas in the majority of such cases (in 687 instances out of 711), subjects did not change their device, in other cases, subjects decided to experiment and input other recommendation devices. While it was not incentive-incompatible for subjects to do so (as they were paid for the best round of the game and thus could not lose from experimenting), there was no fundamental reason to continue experimenting (because, from the feedback, they knew that they had already designed the best device). In these 24 instances (i.e., in 3.38% cases), we imputed the DCE data for all rounds of the game after the DCE was found. If we would simply ignore these data points, our statistics for the latter rounds would distort the subjects' performance in the game.



(a) Dynamics of **TopScore**.



(b) Dynamics of **Score**.

Figure 5: Evolution of the average of points-based performance measures, **TopScore** and **Score**, across rounds for each game. Error bars represent 95% confidence intervals.

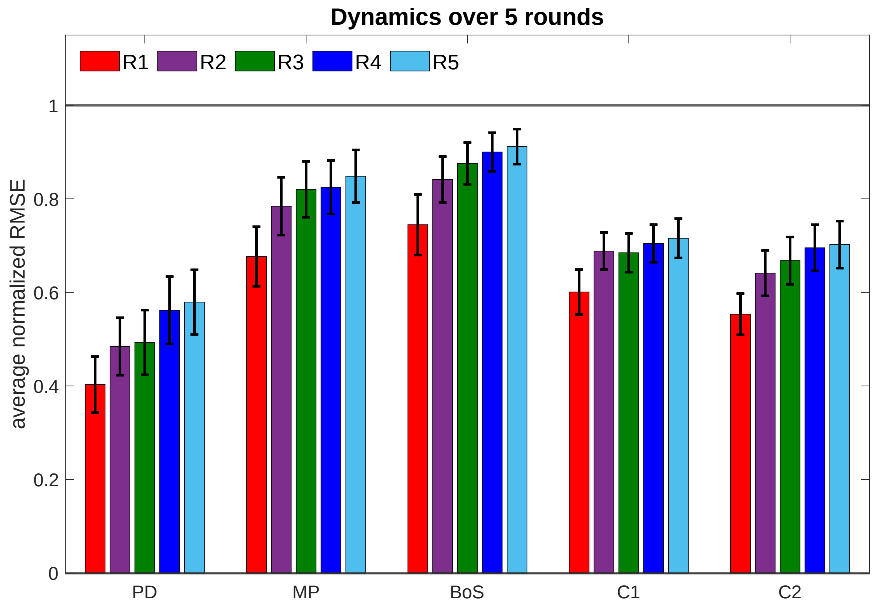


Figure 6: Evolution of the average of distance-based performance measure, N-RMSE, across rounds for each game. Error bars represent 95% confidence intervals.

the fraction of subjects that found the DCE in that round or before. With this interpretation, from Fig. 5a one can see that the majority of subjects find the DCE in both Battle of the Sexes and Matching Pennies quickly: on their first attempt in Battle of the Sexes and by their second attempt in Matching Pennies. Moreover, the rate at which subjects discover the optimum for the first time tapers off in subsequent rounds. This suggests that these games were relatively intuitive for subjects to solve. Instead, for Prisoner's Dilemma and the two versions of Chicken, the rate at which subjects find the DCE is relatively constant across rounds. This is consistent with the idea that solving these games required more experimentation and deliberation on the part of the subjects – evidence that solving these games is relatively more complex.

We perform the same exercise using our distance-based performance measure, N-RMSE, see Fig. 6. We again observe a similar pattern to the points-based measures: performance significantly increases from round 1 but this tapers off relatively quickly: the difference between the N-RMSE in rounds 2 and 5 is insignificant in any game.

Overall, we conclude that the experimental task was something that our subjects

Variable	Average	Description
Order	3.00	order in which the game was played (1-first : 5-last)
StudyYear	2.96	year of study of subject (1-Freshman : 6-PhD Student)
CRTNcorr	2.44	number of correct answers on CRT (0-least : 5-most)
StratSoph	0.73	=0 if answer is 20 for 11-20 game, else = 1
Gender	0.54	=1 if gender is Female, else = 0
Arts	0.03	=1 if major is Arts, else = 0
BiologicalSc	0.06	=1 if major is Biological Sciences, else = 0
Business	0.06	=1 if major is business, else = 0
Design	0.01	=1 if major is Design, else = 0
Economics	0.12	=1 if major is Economics, else = 0
Engineering	0.11	=1 if major is Engineering, else = 0
HealthMed	0.02	=1 if major is Health and Medicine, else = 0
Inform&CS	0.10	=1 if major is Information and Computer Science, else = 0
PhysicalSc	0.03	=1 if major is Physical Sciences, else = 0
SocialSc	0.46	=1 if major is Social Sciences, else = 0

Table 3: List of all potential correlates with performance. Column "Ave" reports their average values (proportions in case of categorical variables).

were able to learn and improve in both through experience and by utilizing the feedback we provided them between rounds. Nonetheless, there are diminishing returns to performance-improvement and so one might expect full convergence to the DCE to require a large number of rounds of experience. This is especially true for games in which only a minority of subjects were able to find the DCE by Round 5 (Prisoner's Dilemma and both versions of Chicken).

5.1.4 Correlate Analysis

We now investigate whether any of the post-experiment information we collected correlates with performance in the experiment. Since the game-order was randomized across subjects, we also check whether this had any impact on performance. Table 3 provides a list of all correlates we consider.

We regress the highest (over five rounds) of each performance measure for each subject in each game on the set of correlates. We also include dummies that control for the exact game that the performance measure corresponds to. The regression

Regressors	**TopScore**	**Score**	**N-RMSE**
PD	0.1094	0.1834**	0.6117***
MP	0.4700***	0.5440***	0.8412***
BoS	0.5303***	0.7148***	0.8519***
C1	−0.0792	0.5874***	0.7395***
C2	0.0396	0.6578***	0.6968***
Order	−0.0036	−0.0034	−0.0068
StudyYear	0.0289	0.0243	0.0099
CRTNcorr	0.0360***	0.0306***	0.0143**
StratSoph	0.0867	0.0532	0.0319
Gender	−0.0227	−0.0043	−0.0048
Arts	−0.1360	−0.1086	−0.0498
BiologicalSc	−0.0965	−0.1391**	−0.0285
Business	0.1548	−0.0588	0.0279
Design	0.1328	0.1458	0.0650
Economics	0.0813	0.0904*	0.0212
Engineering	0.1693**	0.1035*	0.0688**
HealthMed	−0.1258	−0.2793**	−0.0906
Inform & CS	0.0689	0.0055	−0.0053
PhysicalSc	0.0280	0.0667	−0.0283
Diagnostic: adjR^2	0.2833	0.2734	0.2051
* 10%, ** 5%, *** 1% levels of significance (two-sided t-test)			

Table 4: Regressions of the highest (over five rounds) **TopScore**, **Score** and **N-RMSE** on the set of correlates collected in Table 3. Game-specific intercepts are represented by the coefficients on the game dummies. The coefficients on majors are in deviations from the Social Science major that is used as the base category. These coefficients are jointly significantly different from zero (F-test at 5% level) only for the **Score** measure. Each regression uses 500 observations: one observation from each of the five games for each of the 100 subjects.

results are in Table 4.

The findings of our correlate analysis are consistent across three performance measures. First, note that performance on the cognitive reflection test is a significant indicator of success in the experimental task (at the 1% level for both points-based performance measures and at the 5% level for the distance-based performance measure). This is consistent with the fact that our experimental task required substantial contemplative effort rather than intuitive reasoning. In terms of majors, subjects that major in Engineering appear to be better than those that undertake other majors. Perhaps, this is because engineers are more comfortable with the logic of robot-reasoning relative to the broader population.

It is also informative to examine the potential correlates that do not end up exhibiting a significant relationship. First, we see no evidence of gender effects. Second, we do not observe any order effects. This means that performance within a particular game did not appear to be impacted by precisely when subjects designed recommendation devices for that game. Rather, improvements in performance mostly occur with experience *within* a game across rounds, as discussed in Section 5.1.3.

5.2 What Devices Did Subjects Design?

We now explore in more detail the types of recommendation devices that subjects designed in each game. Fig. 7 provides a graphical representation of the average device designed in each round for each game. This is computed by simply taking the average of the designed distribution in each scenario across all players.[16] We again use our amended dataset as described in Section 5.1.3.

Starting with Prisoner's Dilemma, note that the unique correlated equilibrium in this case is the unique Nash equilibrium in pure strategies, i.e., $\mu^*_{RR} = 1$. While subjects are increasing the weight they put on action profile (R, R) round-on-round, convergence to the DCE is slow: by round 5 the average device puts only approximately 50% weight on (R, R). This is because subjects do not give up on the profile that is socially desirable, (B, B), as average weight on this action profile remains somewhat constant from round 2 to round 5. This is consistent with the fact that subjects performed worst in Prisoner's Dilemma in terms of the distance of their designed distributions from the DCE (recall Fig. 4).

For both Matching Pennies and Battle of the Sexes we see that, by round 5, the

[16]Note that, if μ_i is a device designed by subject i, then the average over N subjects, $\overline{\mu} \equiv \sum_{i=1}^{N} \mu_i / N$, is itself a feasible distribution over action profiles $\{R, B\}^2$; that is, $\bar{\mu} \in \Delta(\{R, B\}^2)$.

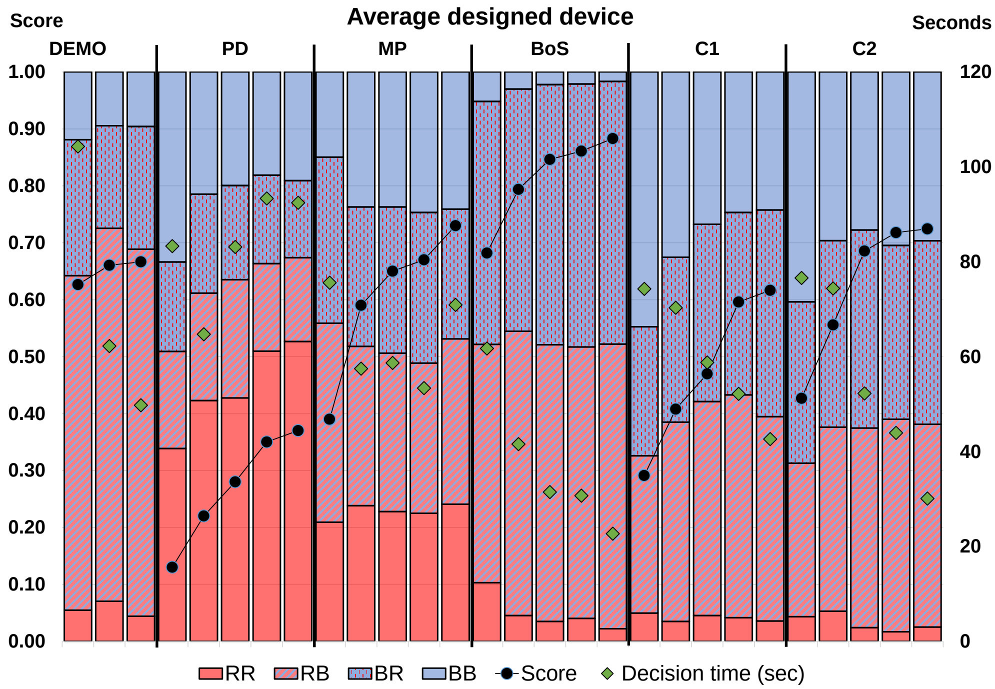


Figure 7: Average designed device for each game across rounds (from round 1 to round 5 reading left to right within each game bin). Black circles indicate the average **Score** (see the left axis). Green diamonds represent the average time spent designing the device in seconds (see the right axis).

average recommendation device is very close to that which subjects were incentivized to implement (1/4 probability on all action profiles for Matching Pennies and 1/2 probability on profiles (R, B) and (B, R) for Battle of the Sexes). This is in line with the performance results for both of these games: subjects consistently implemented the DCE and were close to it when this was not the case, as elucidated by Fig. 4.

Finally, for the two versions of the Chicken game, we see that subjects, on average, were successful in eliminating the extremely undesirable action profile (R, R). However, they did not consistently implement the precise distribution that was optimal.[17] In fact, by round 5, it seems that subjects, on average, *underestimate* the extent to which they can recommend the socially desirable profile in these games.

While inspecting the aggregate data provides insight into the average recommen-

[17]Recall that for Chicken 1, DCE places 1/3 probability on (R, B), (B, R), and (B, B), and, for Chicken 2, DCE places 1/4 probability on (R, B) and (B, R) and 1/2 probability on (B, B).

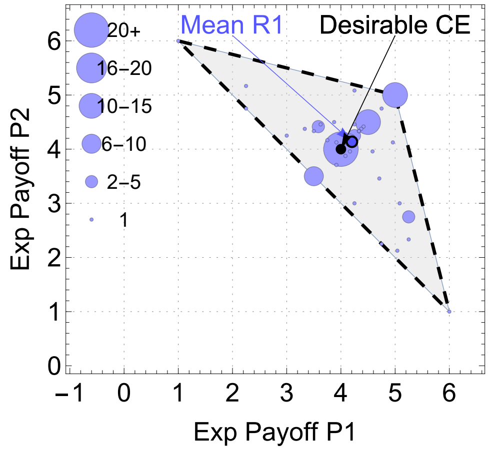
dation behavior of participants, it is also informative to investigate the distribution of recommendation devices designed at the individual level. Doing so allows us to more properly investigate how subjects attempted to solve each game. For each game, we explore the distribution of recommendation devices designed by subjects in round 1 (i.e., their first attempt) and round 5 (i.e., their last attempt or the DCE if achieved in an earlier round). The results are illustrated both in the space of expected payoffs of players and in the distribution space in coordinates $(\mu_{RB}, \mu_{BR}, \mu_{BB})$, cf. Table 1.

5.2.1 Prisoner's Dilemma

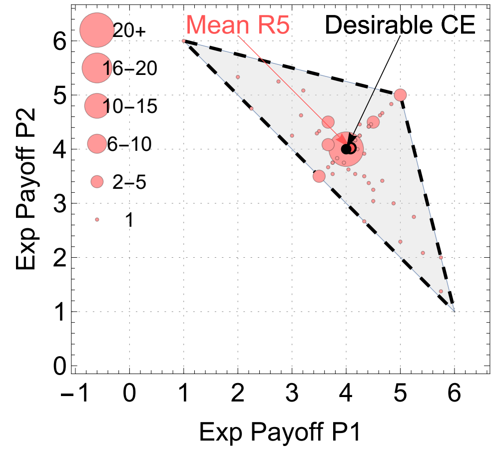
Fig. 8c displays that, in Round 1 of Prisoner's Dilemma, there are a number of salient distributions that subjects try to implement, including both the DCE, (R, R), the socially desirable profile, (B, B), and a 50-50 mix between these two profiles. In general, subjects are putting too much weight on the (B, B), as represented by a distinct number of devices implying expected payoffs in the upper-right quadrant from the DCE in Fig. 8a. Instead, Fig. 8d shows that, by round 5, the modal response of subjects is now clearly the DCE. Nonetheless, the majority of recommendation devices are sub-optimal and such devices appear almost randomly dispersed in the distribution space. Thus, there does not seem to be any consistency in the devices designed by subjects who were unable to find the DCE by round 5. There is still a clear tendency, however, for devices in round 5 to heavily weight the socially-efficient profile, (B, B), as illustrated by the fact that a significant number of devices imply average payoffs that lie on the 45-degree line above the expected-payoff pair implied by the desirable correlated equilibrium (Fig. 8b).

5.2.2 Matching Pennies

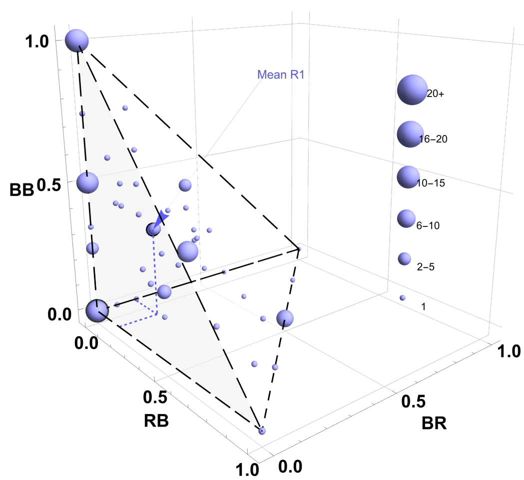
Recall that overall performance in Matching Pennies was high according to all of our performance measures (Fig. 4). Obviously, this would follow from the fact that the subjects successfully designed their recommendation device in this game. Fig. 9d



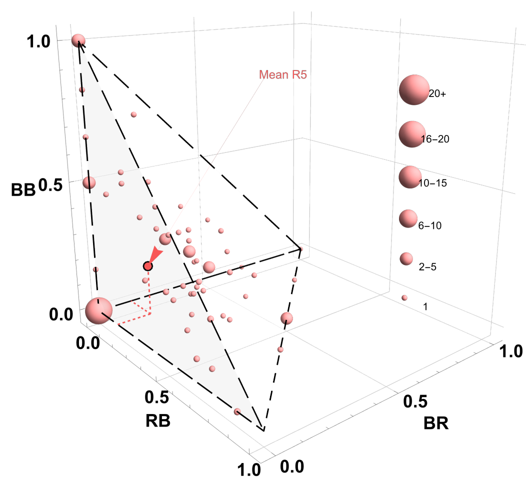
(a) Payoff Space: Round 1



(b) Payoff Space: Round 5



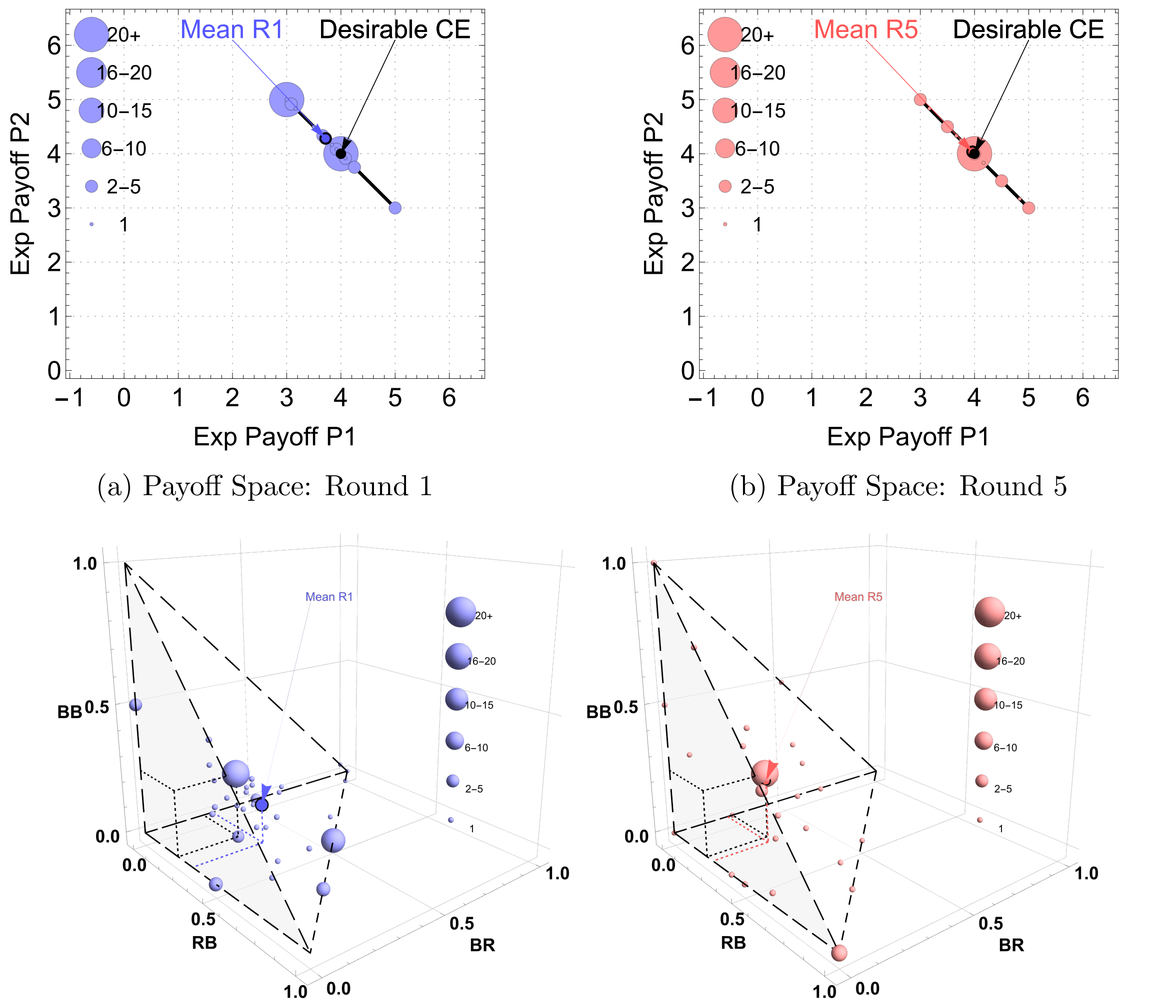
(c) Distribution Space: Round 1



(d) Distribution Space: Round 5

Figure 8: Bubble plots displaying, in both the payoff-space and distribution-space, the distribution of designed devices for **Prisoner's Dilemma** in round 1 vs. round 5.

illustrates this point: by round 5 the clear modal recommendation device is the DCE which coincides with the unique Nash equilibrium in mixed strategies. Moreover, recommendation devices in round 5 that do not constitute correlated equilibria appear to be distributed somewhat randomly around the unique Nash equilibrium, with all of these implying expected payoffs close to those implied by the DCE (Fig. 9b). Thus, subjects appear to recognize the zero-sum nature of Matching Pennies and no additional recommendation devices emerge as focal by the final round.



(a) Payoff Space: Round 1

(b) Payoff Space: Round 5

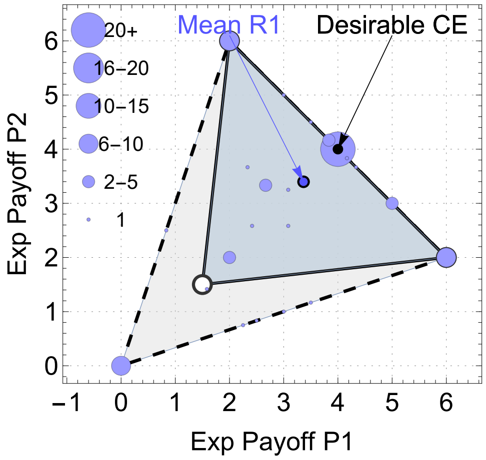
(c) Distribution Space: Round 1

(d) Distribution Space: Round 5

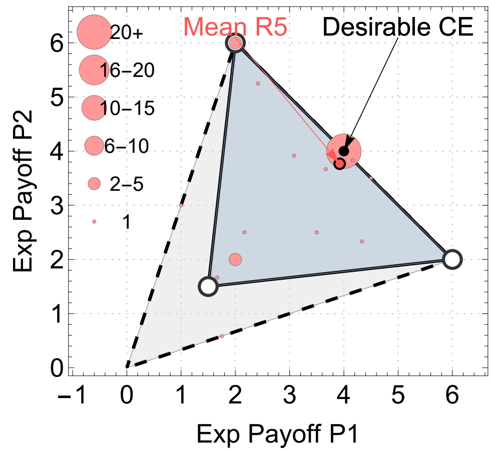
Figure 9: Bubble plots displaying, in both the payoff-space and distribution-space, the distribution of designed devices for **Matching Pennies** in round 1 vs. round 5.

5.2.3 Battle of the Sexes

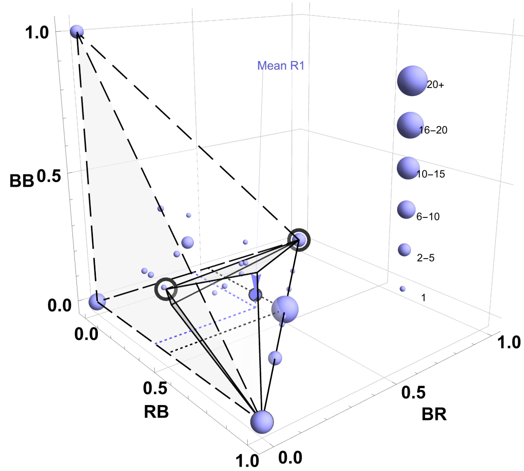
Battle of the Sexes was another game in which subjects performed well in terms of finding the DCE (Fig. 4). From Figs. 10a and 10c, we see that the DCE is the modal response already in round 1, although some subjects implement one of the two asymmetric Nash equilibria with their device. Moreover, already in round 1, there are very few recommendation devices that yield payoffs outside of the set of correlated-equilibrium payoffs. Experience only consolidates this: by round 5 designed



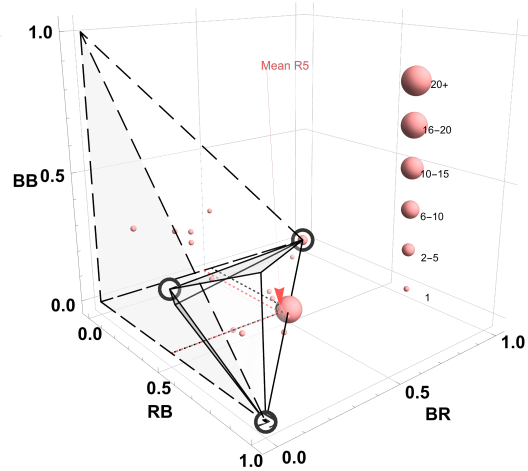
(a) Payoff Space: Round 1



(b) Payoff Space: Round 5



(c) Distribution Space: Round 1



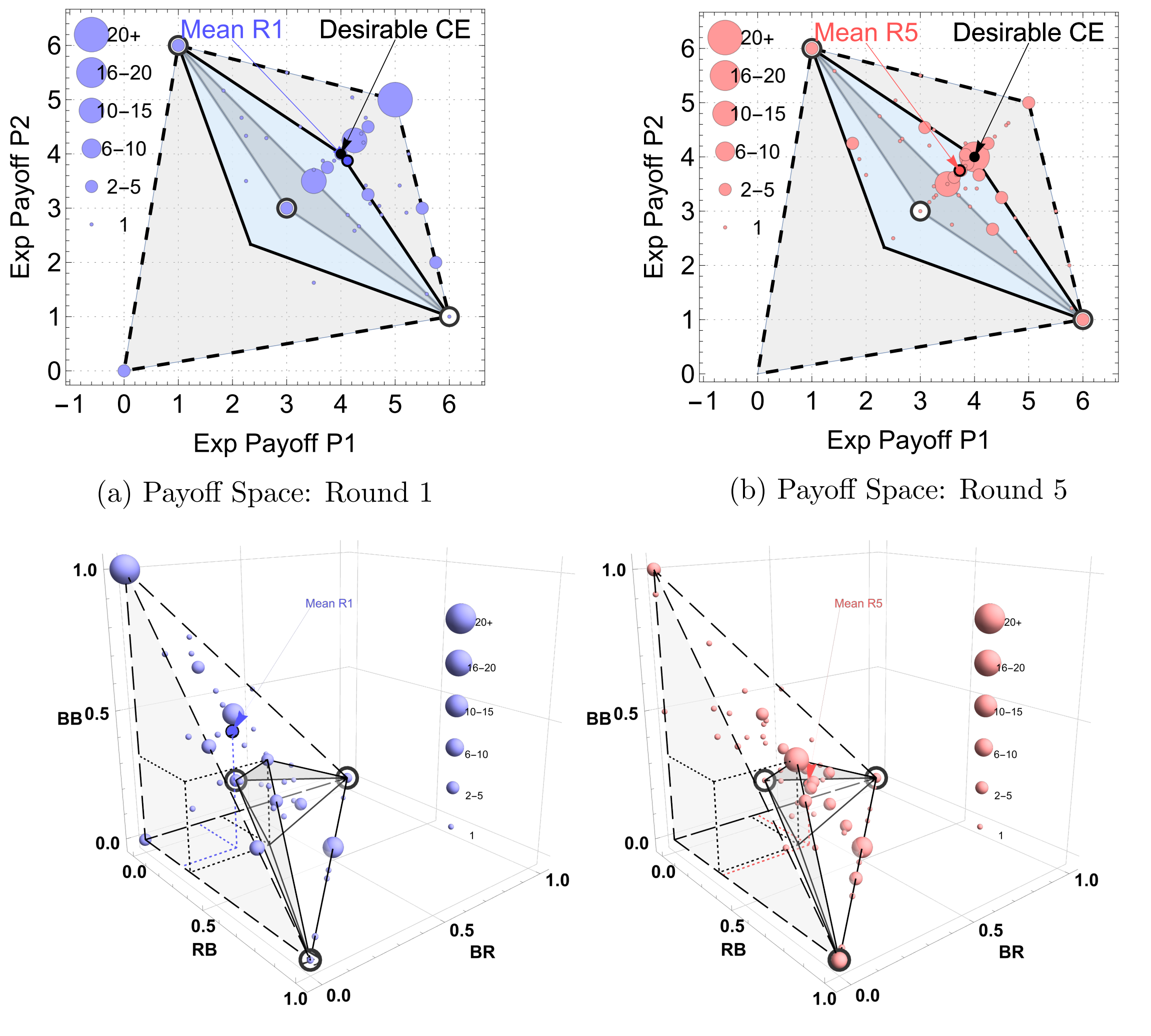
(d) Distribution Space: Round 5

Figure 10: Bubble plots displaying, in both the payoff-space and distribution-space, the distribution of designed devices for **Battle of the Sexes** in round 1 vs. round 5.

distributions are almost entirely in the set of correlated equilibria (Fig. 10d) or the set of correlated equilibrium payoffs (Fig. 10b). The vast majority of these distributions fall on the DCE, as expected from the high aggregate performance in this game.

5.2.4 Chicken, Version 1

Recall that in the Chicken 1 game, the DCE puts zero probability on profile (R, R) and equal probability on the remaining action profiles, including the socially desirable



(a) Payoff Space: Round 1

(b) Payoff Space: Round 5

(c) Distribution Space: Round 1

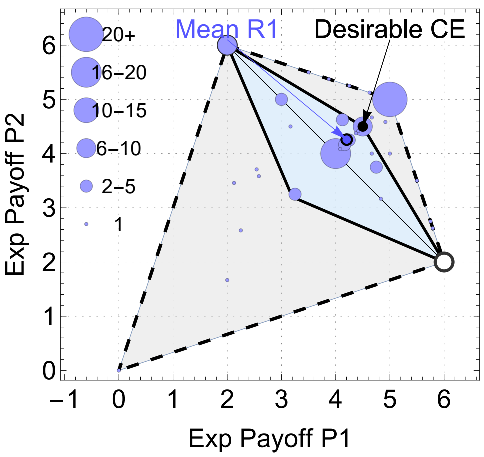
(d) Distribution Space: Round 5

Figure 11: Bubble plots displaying, in both the payoff-space and distribution-space, the distribution of designed devices for **Chicken 1** in round 1 vs. round 5.

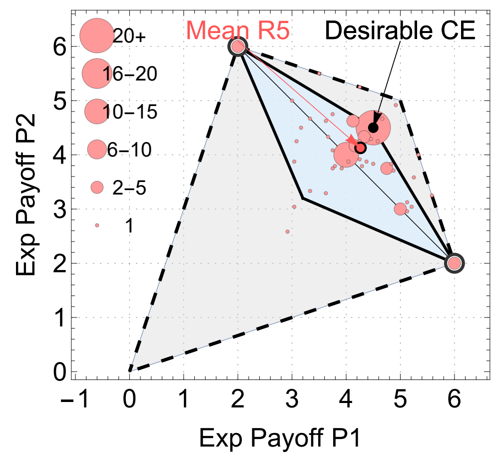
profile, (B, B). From Fig. 11c, one can see that subjects start out optimistic: they put more weight on the socially desirable profile than strategic incentives allow, strongly evidenced by the fact that the recommendation device that puts probability one on (B, B) is the modal device and that many devices imply expected payoffs that are to the upper-right of the DCE payoffs (Fig. 11a). By the final round, however, subjects appear to have learned that they cannot recommend (B, B) too often. Indeed, the majority of recommendation devices fall within the set of correlated equilibria at this point (Fig. 11d). The two most salient recommendation devices at round 5 are the DCE (which is the modal device) and one that mixes 50-50 between the two asymmetric Nash equilibria, (R, B) and (B, R). There are also a number of devices that imply expected payoffs on the 45-degree line between the resulting payoffs from these two salient devices (Fig. 11b). This suggests that subjects understood that partial correlation, in the form of mixing in some (B, B) with the asymmetric action profiles, was optimal. Very few subjects still design devices that over-weight the socially desirable action profile – if anything, the majority of subjects appear to under-estimate the extent to which (B, B) can be recommended which is consistent with the average recommendation device for Chicken 1 in Fig. 7.

5.2.5 Version 2 of Chicken

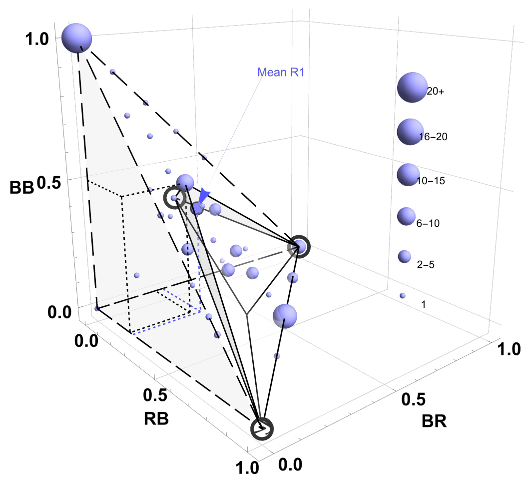
Version 2 of Chicken is similar to version 1 in the sense that the DCE puts zero weight on the inefficient action profile (R, R). The main difference is that subjects can put more weight on the socially desirable profile in this version ($\mu^*_{BB} = 1/2$ rather than 1/3 as in Chicken 1). Examining Fig. 12, we can see that the distribution of designed devices for Chicken 2 is similar to what was found in Chicken 1. In round 1, subjects over-estimate the extent to which they can recommend the socially desirable action profile: the modal device puts probability one on (B, B) (Fig. 12c). However, as was the case in Chicken 1, by round 5 subjects appear to have qualitatively understood that they cannot recommend (B, B) too often. Instead, most recommendation devices now constitute correlated equilibria. The modal device implements the DCE, but the device that mixes 50-50 between the two asymmetric Nash equilibria, (R, B) and (B, R), also appears salient (Fig. 12d). Fig. 12b also shows that there are a number of devices that implement expected payoffs on the 45-degree line between these two salient devices, which again suggests that subjects understood that partial correlation of the actions of the robot-players was required to some extent. Consequently, similar to Chicken 1, the average designed device puts insufficient probability on the socially desirable action profile, as was observed in Fig. 7.



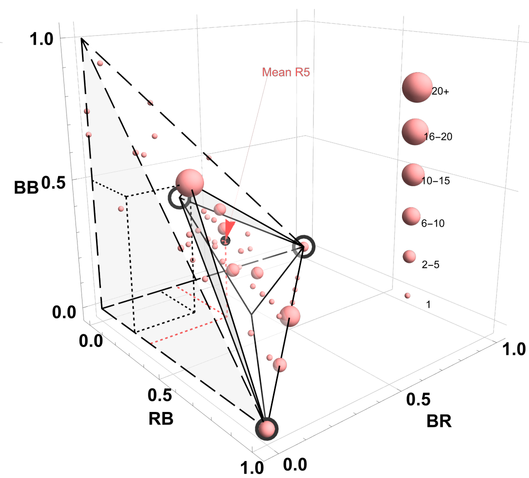
(a) Payoff Space: Round 1



(b) Payoff Space: Round 5



(c) Distribution Space: Round 1



(d) Distribution Space: Round 5

Figure 12: Bubble plots displaying, in both the payoff-space and distribution-space, the distribution of desgined devices for **Chicken 2** in round 1 vs. round 5.

5.3 Discussion

Our analysis suggests subjects succeeded more in some games relative to others, as evidenced by significant differences in performance measures across games (Fig. 4). In particular, subjects were very successful at finding the desirable correlated equilibrium in Battle of the Sexes and Matching Pennies but struggled to do so in Prisoner's Dilemma and both versions of Chicken (although subjects were able to find a correlated equilibrium in both Chicken 1 and Chicken 2). It is interesting to identify whether there are defining features of these particular games and their corresponding correlated equilibria that have led to these observations.

Recall that the DCE in each game required different aspects of equilibrium reasoning in terms of (a) the necessity of randomization, (b) the degree of correlation required, and (c) the alignment of strategic incentives with socially desirable outcomes, see Table 2. We now discuss our results in light of these different aspects.

5.3.1 The Necessity of Randomization

We find no evidence that subjects are not able to design a device that generates random recommendations. Performance is maximized for Matching Pennies and Battle of Sexes, both of which require randomization. Moreover, the only game for which the DCE does not require randomization is Prisoner's Dilemma which is the game for which aggregate performance was lowest.[18]

We would not, however, necessarily conclude that it is the absence of randomization that makes it difficult to grasp the equilibrium reasoning required for Prisoner's Dilemma. Indeed, one might wonder whether subjects were averse to designing a device that made pure-strategy recommendations. We see no evidence of this – there were a number of salient pure-strategy recommendation devices that subjects designed including putting probability one on socially desirable action profiles in Prisoner's Dilemma and Chicken games (Fig. 8a, Fig. 11a, Fig. 12a) and on asymmetric pure-strategy Nash equilibria in Battle of the Sexes and Chicken games (Fig. 10, Fig. 11, Fig. 12). We also provided examples of devices that recommended only a single action profile in the instructions in order to make salient that such devices could constitute optimal devices.

[18]This is consistent with the work of Romero and Rosokha (2023) that suggests individuals are able to randomize by playing mixed strategies in repeated Prisoner's Dilemma when provided with an interface to do so.

5.3.2 The Degree of Correlation

Our results suggest that the degree of correlation that the optimal recommendation device needed to exhibit had a relatively minor impact on success. Subjects had the most success with both the game that required perfect correlation (Battle of the Sexes) and the game that required no correlation (Matching Pennies).

There is some evidence that subjects found it difficult to implement the partial correlation of actions required to solve each version of the Chicken game. However, this appears to be an issue of nuance: subjects didn't find the precise quantification of correlation required. They did appear to understand qualitatively that they needed to correlate the actions of the robot-players using private signals as discussed in Section 5.2.4 and Section 5.2.5.

Again, the game that subjects had least success in was the Prisoner's Dilemma, in which the desirable correlated equilibrium requires no correlation in the sense that it constitutes a pure-strategy Nash equilibrium. However, it does not seem to be the absence of correlation that leads subjects to low performance in the Prisoner's Dilemma. As already emphasized, subjects perform well in Matching Pennies, which also does not require robot-player behavior to be correlated.

5.3.3 The Alignment of Strategic Incentives

The final aspect of equilibrium reasoning we discuss is the extent to which the strategic incentives in each game are aligned with socially desirable outcomes, in the sense of the social objective provided in (1). Recall that the socially desirable outcome is the unconstrained solution to (1), which essentially reduces to finding the distribution over action profiles that is fair and Pareto efficient. As summarized in Table 2, the extent to which strategic incentives are aligned with the socially desirable outcome varies significantly across the games considered. We believe that this variation has substantial explanatory power for heterogeneity in design-success across games.

In order to facilitate this discussion, we provide a quantitative measure of the

	μ^{SD}	μ^*	N-RMSE($\mu^{\text{SD}}; \mu^*$)
PD	$(0,0,0,1)$	$(1,0,0,0)$	0
MP	$\left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)$	$\left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)$	1
BoS	$\left(0,\frac{1}{2},\frac{1}{2},0\right)$	$\left(0,\frac{1}{2},\frac{1}{2},0\right)$	1
C1	$(0,0,0,1)$	$\left(0,\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$	0.293
C2	$(0,0,0,1)$	$\left(0,\frac{1}{4},\frac{1}{4},\frac{1}{2}\right)$	0.478

Table 5: The socially desirable outcome, μ^{SD}, the DCE, μ^*, and the normalized distance (based on the Root Mean Square Error) between them for each game.

alignment of strategic incentives. To do this, we measure how far the desirable correlated equilibrium, μ^*, is from the socially desirable outcome, μ^{SD}, for each game using our distance measure, N-RMSE. This information is summarized in Table 5.

According to Table 5, in Matching Pennies and Battle of the Sexes strategic incentives are perfectly aligned with the socially desirable outcome as μ^* and μ^{SD} coincide for these games. Instead, strategic incentives and the socially desirable outcome are maximally unaligned for Prisoner's Dilemma as μ^* is maximally distant from μ^{SD} in this game. Finally, for both versions of Chicken, strategic incentives and the socially desirable outcome are partially aligned as N-RMSE(μ^*, μ^{SD}) is strictly greater than zero for these games. Specifically, μ^* is closer to μ^{SD} for C2 relative to C1. This implies the following order on the extent to which strategic incentives and the socially desirable outcome are aligned across games:

$$\mathbf{MP} \sim \mathbf{BoS} \succ \mathbf{C2} \succ \mathbf{C1} \succ \mathbf{PD}. \tag{2}$$

We now interpret our results in light of the order in (2). First, note that Matching Pennies and Battle of the Sexes are the games for which subjects performed the best, and also happen to be the games in which strategic incentives and the socially desirable outcome are perfectly aligned.

Next, we have the two versions of Chicken in which strategic incentives are only weakly aligned with the socially desirable outcome. While subjects were not particularly successful in implementing the DCE for these games, they were able to consistently find correlated equilibria that were close to the optimum (evidenced by high **Score** and **N-RMSE** in Fig. 4). Moreover, on their final attempt, it does not appear that subjects over-estimate the extent to which they can recommend (B, B) but, rather, tend to under-weight this action profile as discussed in Sections 5.2.4 and 5.2.5. This suggests that subjects understand, at least qualitatively, the role of strategic incentives in the Chicken games, which may be due to the fact these strategic incentives are weakly consistent with the socially desirable outcome. Moreover, by the relation in (2), strategic incentives are more closely aligned with the socially desirable outcome in Chicken 2 relative to Chicken 1. From Fig. 4, our subjects were more likely to find the DCE in Chicken 2 relative to Chicken 1 (approximately 10% more subjects for Chicken 2) and achieved a higher **Score**, although these differences are not statistically significant.

The game for which strategic incentives are completely at odds with the socially desirable outcome is Prisoner's Dilemma. This is because the DCE puts zero weight on the socially desirable profile (B, B) in this game. Subjects struggled substantially with Prisoner's Dilemma: the majority don't find the DCE and their designed devices are relatively distant from the optimum (Fig. 4). This seems to be primarily driven by the conflict between strategic incentives and the social objective in Prisoner's Dilemma. Indeed, subjects consistently design devices that put a positive probability on (B, B) through to the last round (Figs. 8b and 8d) and the average weight placed on this action profile does not fall significantly with experience over rounds (Fig. 7).

6 Conclusion

Our main contribution in this paper has been to provide a methodology for testing how people reason about equilibrium while abstracting from issues that are usually prevalent in such situations, such as strategic uncertainty and social preferences. We do this by having subjects design recommendation devices for Bayesian, expected-utility maximizing robot-players that only follow recommendations if it is a best response to do so.

Overall, subjects in our study found success in the task of designing recommendation devices for different games, under the objective of implementing correlated equilibria that are Pareto efficient and fair. As discussed in Section 5.1.3, their performance significantly improved across rounds, and our analysis in Section 5.2 suggests that the modal device designed by subjects in every game implements the desirable correlated equilibrium.

There is, however, significant heterogeneity in performance across games. We find that subjects perform well in designing recommendation devices for both Battle of the Sexes and Matching Pennies: the majority of subjects find the desirable correlated equilibria in these games, and those that don't still design devices that are close to optimal. Subjects had less success finding the desirable correlated equilibrium in Prisoner's Dilemma and in the two versions of the Chicken game. Nonetheless, they do successfully find a correlated equilibrium in Chicken even if they don't solve the quantitative problem of finding the correlated equilibrium that is precisely optimal. By contrast, for Prisoner's Dilemma, subjects struggle significantly with the qualitative aspects of equilibrium reasoning implied by the strategic incentives in this game: those that do not solve Prisoner's Dilemma tend not to give up on the possibility that the socially desirable profile can be recommended with strictly positive probability. We believe that these findings are consistent with the idea that people find it hard to grapple with strategic incentives that run counter-productive to achieving socially desirable outcomes, as discussed extensively in Section 5.3.

Our findings suggest that people bring a cooperative mindset into reasoning about equilibrium governed by non-cooperative incentives, at least in situations in which the focus is on achieving socially desirable outcomes. It appears to be difficult for subjects to reconcile the fact that strategic incentives can be so at odds with desirable outcomes, and they do not come to the realization that there is no device that utilizes the efficient profile in Prisoner's Dilemma. Interestingly, this holds even though the strategic incentives in Prisoner's Dilemma are extremely strong: each robot-player, of course, has a *strictly dominant strategy* to choose Red. One might expect that, as strategic incentives become weaker, it will become even more difficult for subjects to recognize situations in which socially desirable outcomes cannot be achieved. It would be interesting to investigate, however, whether this cooperative mindset can be manipulated by changing the design objective.

Our framework is sufficiently flexible so as to investigate how individuals fare in finding correlated equilibria in any two-player game. In this paper, we have focused on a set of canonical games that are symmetric. It would be interesting, however, to explore other games, including those that have asymmetric payoffs. It would also be interesting to explore whether individuals can successfully coordinate the behavior of more than two players. Finally, it would be interesting to understand whether individuals reason differently when designing recommendation devices for other humans, rather than for robots. These ideas will be explored in future research.

References

Anbarci, N., N. Feltovich, and M. Y. Gürdal (2018): "Payoff Inequity Reduces the Effectiveness of Correlated-Equilibrium Recommendations," *European Economic Review*, 108, 172–190.

Arad, A. and A. Rubinstein (2012): "The 11–20 Money Request Game: A Level-k Reasoning Study," *American Economic Review*, 102, 3561–3573.

Aristidou, A., G. Coricelli, and A. Vostroknutov (2019): "Incentives or persuasion? An experimental investigation," Available at: https://doi.org/10.26481/umagsb.2019012.

Au, P. H. and K. K. Li (2018): "Bayesian Persuasion and Reciprocity: Theory and Experiment," Available at SSRN 3191203.

Aumann, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

——— (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 1–18.

Bone, J., M. Drouvelis, and I. Ray (2013): "Coordination in 2 x 2 Games by Following Recommendations from Correlated Equilibria," *Working Paper*.

Cason, T. N. and T. Sharma (2007): "Recommended Play and Correlated Equilibria: An Experimental Study," *Economic Theory*, 33, 11–27.

Cason, T. N., T. Sharma, and R. Vadovic (2022): "Correlated Equilibria and Forecasts based on Naïve Play in Hawk-Dove Games," *Available at SSRN 4174461*.

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Chen, Y. and J. O. Ledyard (2010): "Mechanism Design Experiments," in *Behavioural and Experimental Economics*, ed. by S. N. Durlauf and L. E. Blume, Palgrave Macmillan, 191–205.

Duffy, J. and N. Feltovich (2010): "Correlated Equilibria, Good and Bad: An Experimental Study," *International Economic Review*, 51, 701–721.

Duffy, J., E. K. Lai, and W. Lim (2017): "Coordination Via Correlation: An Experimental Study," *Economic Theory*, 64, 265–304.

Enke, B. and F. Zimmermann (2019): "Correlation Neglect in Belief Formation," *The Review of Economic Studies*, 86, 313–332.

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

Fréchette, G. R., A. Lizzeri, and J. Perego (2022): "Rules and Commitment in Communication: An Experimental Analysis," *Econometrica*, 90, 2283–2318.

Frederick, S. (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19, 25–42.

Friedman, D., J. P. Rabanal, O. A. Rud, and S. Zhao (2022): "On the Empirical Relevance of Correlated Equilibrium," *Journal of Economic Theory*, 205, 105531.

Georgalos, K., I. Ray, and S. SenGupta (2020): "Nash versus Coarse Correlation," *Experimental Economics*, 23, 1178–1204.

Gintis, H. (2014): *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences-Revised Edition*, Princeton University Press.

Hart, S. and A. Mas-Colell (2000): "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, 68, 1127–1150.

Hossain, T. and R. Okui (2021): "Belief Formation Under Signal Correlation," *Available at SSRN 3218152*.

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Moreno, D. and J. Wooders (1998): "An Experimental Study of Communication and Coordination in Noncooperative Games," *Games and Economic Behavior*, 24, 47–76.

Myerson, R. B. (1991): *Game theory: analysis of conflict*, Harvard University Press.

Nash Jr, J. F. (1950): "Equilibrium Points in N-person Games," *Proceedings of the National Academy of Sciences*, 36, 48–49.

Rawls, J. (1971): *A Theory of Justice: Original Edition*, Harvard University Press.

Romero, J. and Y. Rosokha (2023): "Mixed Strategies in the Indefinitely Repeated Prisoner's Dilemma," *Econometrica (Forthcoming).*

Wu, W. and B. Ye (2021): "Competition in Persuasion: An Experiment," Working Paper.

Ziegler, A. (2022): "Persuading an Audience: Testing Information Design in the Laboratory," Working Paper.

Appendix For Online Publication Only

A Experimental Instructions

Welcome

Welcome to this experiment in economic decision-making. Please read these instructions carefully as they explain how you earn money from the decisions you make. For completing this experiment you are guaranteed the $10 show-up fee. Depending on your actions, you can earn an additional amount as explained below. Your earnings during the experiment will be computed in *points* that will be converted to the real dollars at the end of the experiment at the rate of 1 point = $6, so more points means more money.

Please do not talk with others for the duration of this experiment and silence all mobile devices. If you have any questions, please raise your hand.

Overview

Your job will be to design recommendations to two intelligent robot players on how they should play a game with one another. In this game, the two robot players have to simultaneously choose between two actions, labeled Red and Blue. After making their choices, each robot player receives a payoff in points that depends: (1) on the action color they chose and (2) on the action color that the other player chose.

The payoffs of robot Player 1 and of robot Player 2 are shown in the payoff table of the game. Each cell of the table corresponds to the chosen actions of both players. The payoffs of robot Player 1 are shown in the lower left part of the cell, and the payoffs of robot Player 2 are shown in the upper right part of the cell.

The figure below provides an **example** of the payoff table for a game. In this case, if, say, Player 1 (P1) plays Blue and Player 2 (P2) plays Red, then the payoff to Player 1 is 2 points and the payoff to Player 2 is 5 points.

P2 / P1	Red	Blue
Red	1 0	3 3
Blue	5 2	1 3

Games will differ in their payoffs to each player. The specific payoffs will be known to both the robot players and you.

The two robot players are not able to communicate with one another. Further, each game will be played for multiple periods but the robot players in each period will always be different (so that their actions will not depend on what was played in the past).

Your task is **to design a device that will make recommendations to both players** about which color action they should choose in each period. These recommendations are separate, that is, each robot player knows only their own recommendation from your device. However, the robot players know the design of your device and use this knowledge to infer as much as possible about the recommendation the other robot player received.

The robot players are not easily influenced and **do not follow recommendations blindly**. They are self-interested, intelligent and take into account that the other robot player is self-interested and intelligent as well. When a particular color action is recommended to a robot, it will investigate whether it is truly worthwhile to follow this recommendation, using all the information from the design of your recommendation device. The robot player will do this by answering two questions: (a) given the device and the recommendation received, what is the percentage chance that the other robot player received each of their possible recommendations?; and (b) given that the other robot player follows their recommendation, do I (the robot) earn a higher payoff by following the recommendation given to me or by choosing the other color action?

You will design this recommendation device for **five** games that will differ in their payoffs to the two players. You will design a separate device for each game. This will happen sequentially. This means that you will first design the device for one game, then, once completed, for the next game, and so on. For each game, there will be **five** rounds during which you can update your recommendation device to try to increase the amount of money that you take home with you today.

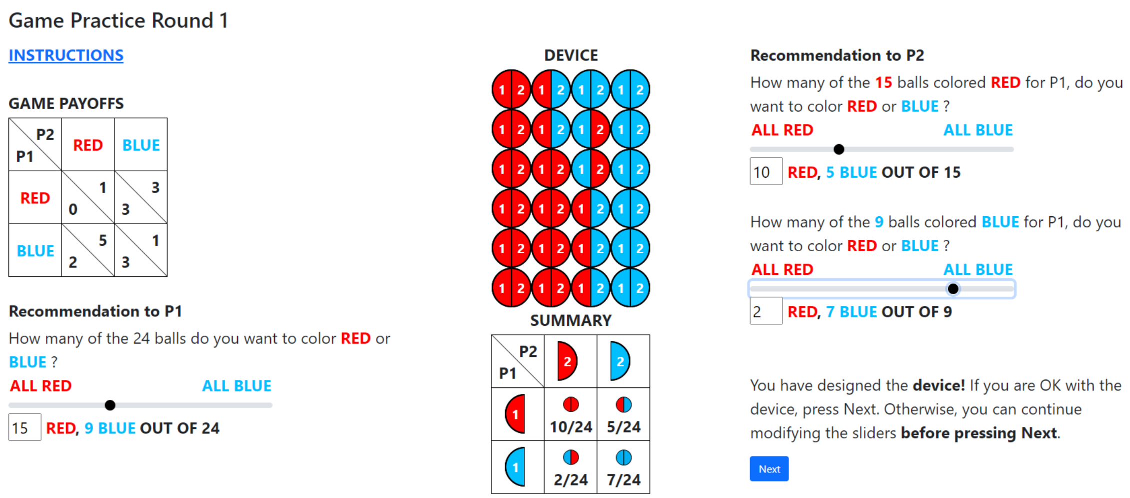
After reading these instructions, you will be asked several control questions. Next, you will play three practice rounds. Right after the practice rounds, the main part of the experiment (5 games with 5 rounds each) will start. After the main part, you will be asked to answer several quiz questions and fill out a short survey. Following completion of the survey, your participation in the study is over and you will be paid your earnings.

Designing the recommendation device

For each game, your task is to design a device that makes separate recommendations to Player 1 and to Player 2 as to which color action to choose, either Red or Blue. With your device, the game will be played multiple periods. In each period, two intelligent robot players will participate in the game, the row robot "Player 1" and the column robot "Player 2".

Your recommendation device will make use of a container filled with 24 balls. Each ball can be split into two halves: one half is labeled **1** and the other half is labeled **2**. Your task is **to select a color for each half of each ball**, Red or Blue. In other words, you will fill the container with 24 balls of four types, RR, RB, BR, and BB, where the first letter corresponds to the half labeled **1** and the second letter to the half labeled **2**.

The picture below shows an example of the screen where you design a device. You first choose the number of balls (out of 24) for which you want to color the half labeled **1** Red or Blue. This choice is in the lower left part of your decision screen, labeled **Recommendation to P1**. Then, of those balls for which you colored the half labeled **1** Red, you will chose the number of balls for which you want to color the half labeled **2** Red or Blue. Similarly, out of those balls for which the half labeled **1** were colored Blue, you will chose the number of balls for which you want to color the half labeled **2** Red or Blue. These two choices are in the upper right part of your decision screen called **Recommendation to P2**. You can color the balls using either a slider or by entering the exact number of Red balls in the box below the slider.



After you have made all selections, your device with the colored balls will be complete. Below the device, the resulting **Summary** table will appear showing the total number of balls of each type. You may wish to consult this summary table to better comprehend how many balls of each type your device contains.

NOTE: you can achieve any desired device! For this, you do the following:

1. Count the number of balls in your desired device where the half labeled **1** is Red. Insert this number as the **Recommendation to P1**.
2. Count the number of RR balls in the desired device, and insert this number in the upper right part of the screen, under **Recommendation to P2**.
3. Count the number of BR balls in the desired device, and insert this number in the middle right part of the screen, under **Recommendation to P2**.

In the previous example, the device has 15 balls where the left part is Red: you can count them inside the container, and you also can find it as a sum of 10 RR and 5 RB balls in the second row of the Summary table. This number 15 is entered in the lower left part of the screen as the **Recommendation to P1**. There are 10 RR balls, and 10 is entered in the upper right as the **Recommendation to P2**. Finally, there are 2 BR balls, and 2 is entered in the middle right part as the **Recommendation to P2**.

You can adjust your device as many times as you want, until you click on the **Next** button, at which point your device is finalized for the round.

After you design the device, it will be used to provide recommendations for many periods of the game. Each period, the game will be played by a new pair of robot players. One ball will be drawn **at random** and will be split in two halves. The half labeled **1** will be given to Player 1 and the half labeled **2** will be given to Player 2. The color of the piece of the ball given to each player represents the recommended color action for each player.

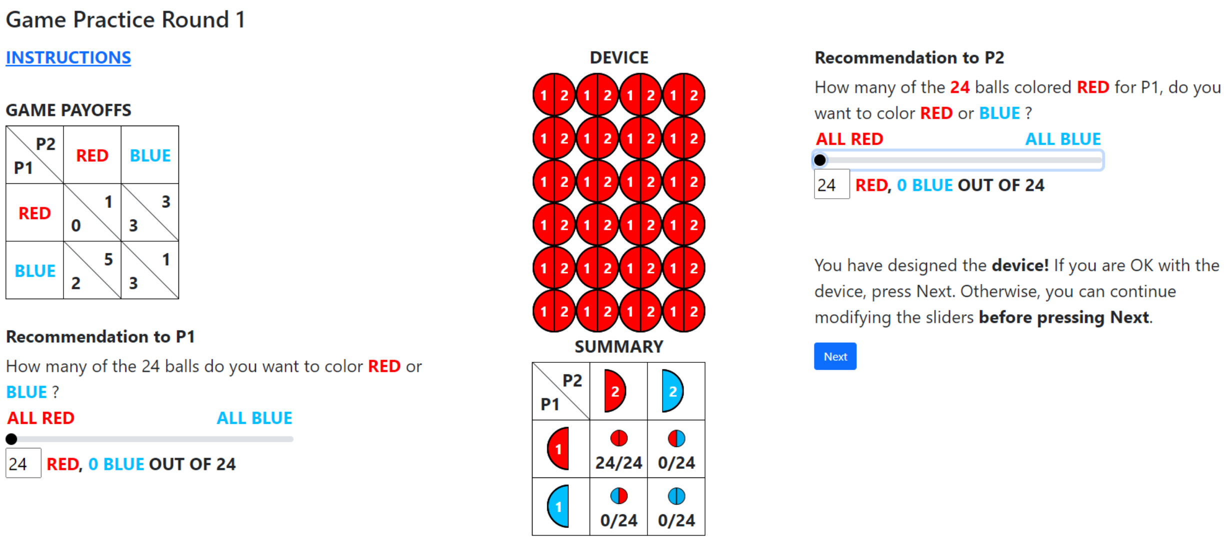
The robot players **know the design of your device**, that is, the proportion of colored balls of each type in the container. They use this knowledge in deciding whether or not to follow the recommendation they received. If they can do better by not following the recommendation of your device, then they will not follow the recommendation of your device. Otherwise, they will follow it.

Examples of Recommendation Devices and Robots' Reasoning

We discuss three examples for the game we introduced earlier:

P1 \ P2	Red	Blue
Red	0 \ 1	3 \ 3
Blue	2 \ 5	3 \ 1

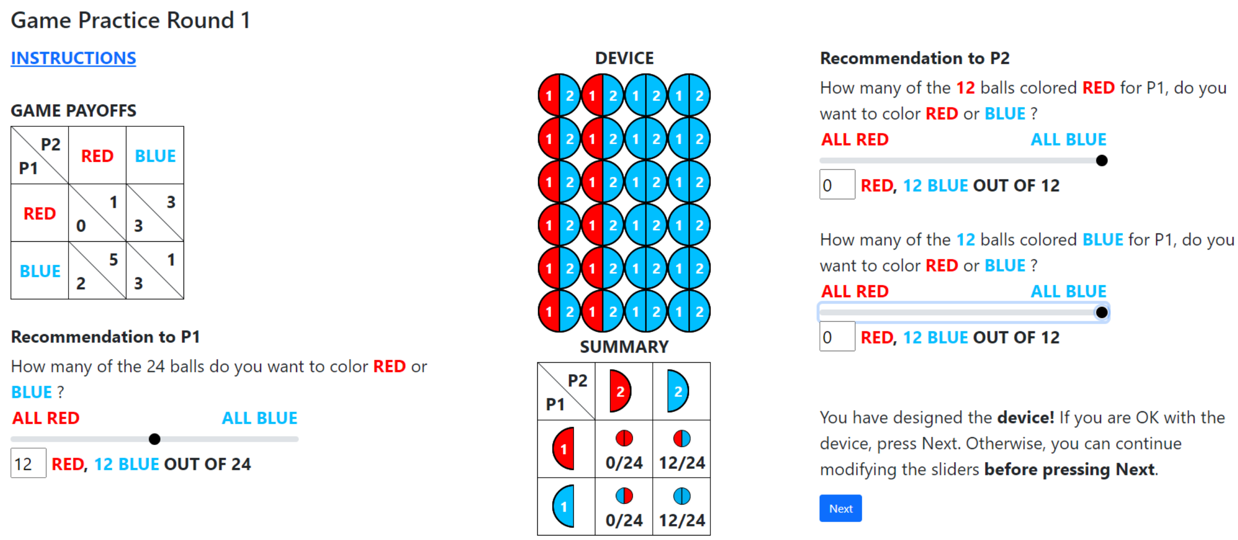
Example 1. Suppose you color all balls RR (i.e., *both halves* of all 24 balls are Red), so that your resulting device looks as follows.



This device always recommends that Player 1 chooses Red and that Player 2 also chooses Red. Even though Player 1 does not see the recommendation given to Player 2, by the design of the device, Player 1 knows that Player 2 is recommended to choose Red. As such, Player 1 compares following the recommendation and receiving a payoff of 0 points (Player 1's payoff when both players play Red) to not following the recommendation, choosing Blue instead and receiving a payoff of 2 points (Player 1's payoff when Player 1 chooses Blue and Player 2 chooses Red). Since a payoff of 0 from playing Red is smaller than 2 from playing Blue, Player 1 does not follow the recommendation of Red. (Player 2 does not follow the recommendation as well. Player 2 infers that Player 1 was recommended Red, and prefers to play Blue.)

In Example 1, Player 1 knew exactly what is recommended to Player 2 and Player 2 knew exactly what is recommended to Player 1. More generally, this will not be the case.

Example 2. Suppose that you color 12 balls as RB and 12 balls as BB so that your device is as shown below:

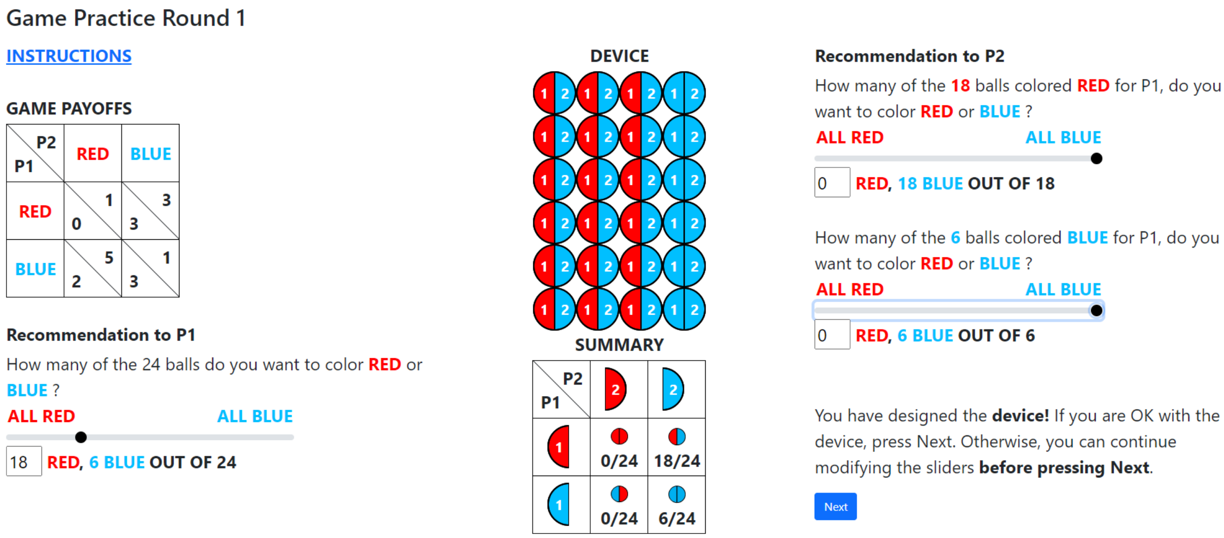


Essentially, Player 2 is always recommended Blue, and Player 1 is recommended Red or Blue with a chance of 50% each. While Player 1 knows that Player 2 is always recommended Blue, Player 2 does not know the recommendation given to Player 1. Player 2 only knows that Player 1 gets a recommendation of Red with a chance of 50% and gets a recommendation of Blue with a chance of 50%.

With this knowledge, Player 2 does not follow the recommendation of Blue. This is because, if Player 2 chooses Blue, then Player 2 receives 3 points when Player 1 chooses Red (which occurs 50% of the time) and 1 point when Player 1 chooses Blue (which occurs 50% of the time). This gives on average $0.5 \times 3 + 0.5 \times 1 = 2$ points. If Player 2 instead chooses Red, Player 2 receives 1 point when Player 1 chooses Red and 5 points when Player 1 chooses Blue, which is on average $0.5 \times 1 + 0.5 \times 5 = 3$ points. Hence, the average payoff from not following the recommendation and choosing Red for Player 2 is higher than following the recommendation of Blue.

This illustrates the calculations that the self-interested, intelligent robots make in deciding whether or not to follow the recommendations of your device. Note that the ultimate reason Player 2 does not follow the recommendation of Blue in this example, is because Player 1 is recommended to choose Blue too often (thereby incentivizing Player 2 to play Red and grab the high payoff of 5). Just because Player 2 does not follow the Blue recommendation does not necessarily mean Player 2 should be recommended to play Red!

Example 3. Consider the following device, where 18 balls are colored as RB and 6 balls are colored as BB.



Compared with Example 2, Player 2 is again always recommended Blue, but Player 1 is now more often recommended to play Red than Blue.

Consider the game from the perspective of Player 2. If Player 2 follows the recommendation to play Blue, and Player 1 follows its recommendations, then Player 2 will receive a payoff of 3 in 75% of all periods (that is when Player 1 is recommended to play Red since $18/24 = .75$ balls have the half labeled **1** colored Red) and a payoff of 1 in 25% of all periods (that is when Player 1 is recommended to play Blue since 6/24=.25 balls have the half labeled **1** colored Blue). Player 2 thus expects to get 2.5 on average by following the recommendation (as $0.75\times3+0.25\times1 = 2.5$). Not following the recommendation and playing Red, will bring a payoff of 1 in 75% of all periods and a payoff of 5 in 25% all periods i.e., 2 on average (as $0.75 \times 1 + 0.25 \times 5 = 2$). It is now better for Player 2 to follow the recommendation to play Blue.

We can also consider the game from the perspective of Player 1 who knows that Player 2 was recommended Blue. Player 1 is indifferent between playing Red and Blue, as the payoff will be 3 in either case (that is, when Player 2 plays Blue and Player 1 plays either Red or Blue). Thus, Player 1 will follow the recommendation Red, if received, and the recommendation Blue, if received. In this example, <u>all recommendations</u> from this device will always be followed.

Your points for a round

For each game, you will participate in 5 rounds. At the beginning of the first round, you will have to design the device. At the end of each round, you will receive feedback on how well your device performed. At the start of each subsequent round, you will have the opportunity to re-design the device.

Each round will give you a certain number of points, which convert into money earnings at a fixed rate. We now describe how this amount of points is determined.

As stated, each round consists of robots playing the game for a **very large** number of periods. In each period, a ball is drawn at random, which represents the recommendation made to each robot player. The robot players decide whether or not to follow their recommendations and play the game once. Then the ball is placed back into the container, and a new period begins. A ball is drawn again for two new robot players, and so on. The number of periods is so large that any particular ball in the container will eventually be drawn at some point.

If in *at least in one period*, one of the robot players did not follow your recommendation, then you will not receive any points in the round. That is, for you to receive points, the robot players must *always* (that is, in all periods) follow the recommendations that your device makes to them. In this case, your **point earnings** are equal to **the smallest of the two average payoffs**, the average payoff earned by player 1 and the average payoff earned by player 2.

We return to our game and provide some examples of payoff calculations.

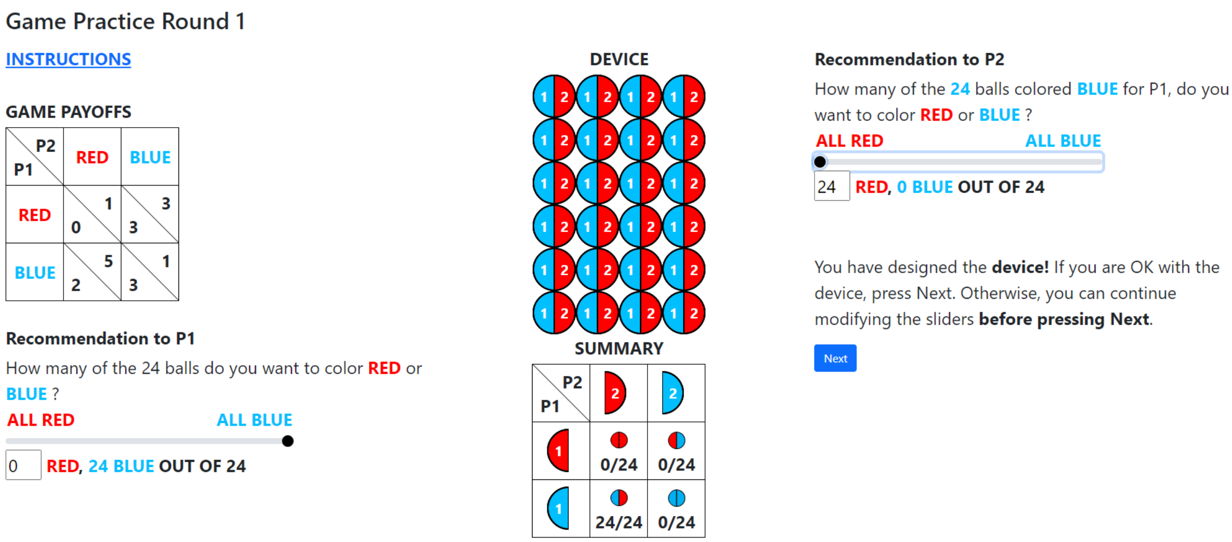
P1 \ P2	Red	Blue
Red	0, 1	3, 3
Blue	2, 5	3, 1

Example 1. If you color all balls RR, then neither Player 1 nor Player 2 would follow the recommendations from this device. You earn **no points**.

Example 2. If you color 12 balls as RB and 12 balls as BB, then Player 2 would not follow the recommendations, as previously discussed. You earn **no points**.

Example 3. If you color 18 balls as RB and 6 balls as BB, then players would follow your recommendations. Player 1 gets 3 points always with this device, and player 2 gets 2.5 points on average, as we calculated. The minimum of 3 and 2.5 is 2.5. Thus, **you earn 2.5 points**.

Example 4. If you color all balls BR, you get this device.



Player 1 knows that Player 2 is recommended Red and follows the recommendation Blue (as the payoff of 2 is not less than the payoff of 0 from not following this recommendation). Player 2 knows that Player 1 is recommended Blue and follows the recommendation Red (as the payoff of 5 is not less than the payoff of 1 from not following this recommendation). Hence, you earn additional points with this device. The minimum average payoff of 2 and 5 is 2. Thus, **you earn 2 points**.

Feedback and updating the device

Between rounds, you will have a chance to update the design of your device (that is, re-color the balls in the container of 24 balls).

After each round, you will receive *feedback* consisting of two parts. Above all, you will learn whether the recommendations produced by your device were always followed by the robot players or not.

Then, in the case the robot players always followed the recommendations of your device, you will also learn whether it is possible to improve your device to receive more points or, instead, the maximum possible payoff has been achieved. In the case where the robot players did not always follow the recommendations, you will see an example of a one-period realization (i.e., one ball drawn from your device) when the recommendations were not followed and why one player chose not to follow the recommendation.

Your total payoff for the experiment

At the end of the experiment, the computer will choose one of the five games at random and you will receive the number of points you earned in **the round for that game with the highest payoff**. This amount will be converted into dollars at the rate of 1 point equal to $6. You will be paid only if you complete all rounds of the experiment, the quiz, and the short survey.

This means that you will definitely earn the show-up fee of $10 in this experiment. If, in addition, your maximum round earnings in the randomly selected game is, for instance, 3 points, then you will receive a payment of $3 \times \$6 = \$18$ (in addition to your guaranteed earnings of $10).

To summarize, the payoff structure of this experiment implies:

- you will definitely earn $10 in this experiment
- you want to do your best in every game to earn an additional amount of points/-money.
- you want to design a recommendation device where both players always follow the recommendations
- you can improve your payoff by improving your recommendation device in a given round (if the device is not yet giving you the highest possible payoff)

- you are encouraged to experiment between rounds by changing the color of the balls in order to try to improve your payoff.

B Experiment Questions

Pre-Experiment Quiz

Question 1: Suppose that the game that the robot players participate in has the following specific payoffs:

P1 \ P2	Red	Blue
Red	1 0	3 3
Blue	5 2	1 3

If Player 1 chooses Blue and Player 2 chooses Blue, then:

(a) The payoff to Player 1 is 3 and the payoff to Player 2 is 3.

(b) The payoff to Player 1 is 5 and the payoff to Player 2 is 2.

(c) The payoff to Player 1 is 3 and the payoff to Player 2 is 1. [*Correct*]

(d) The payoff to Player 1 is 1 and the payoff to Player 2 is 3.

Feedback: "The payoffs of two players are in the cell corresponding to their actions (Red or Blue). The payoff of player 1 is in the lower left part of the cell and the payoff of player 2 is in the upper right part of the cell."

Question 2: Consider the same game

P1 \ P2	Red	Blue
Red	1 0	3 3
Blue	5 2	1 3

Suppose that Player 2 thinks that Player 1 is going to choose Blue. Then,

(a) It is in Player 2's best interest to choose Blue.

(b) It is in Player 2's best interest to choose Red. [*Correct*]

(c) It doesn't matter whether Player 2 chooses Blue or Red.

(d) None of the above are true.

<u>Feedback:</u> "Player 2 thinks that player 1 plays Blue and thus look at the bottom row of the table. Then, player 2 compares his/her payoff from playing Red (that is 5) with his/her payoff from playing Blue (that is 1). Player 2 will choose the action leading to the maximum of these two payoffs."

Question 3: Consider the same game

P1 \ P2	Red	Blue
Red	0 \ 1	3 \ 3
Blue	2 \ 5	3 \ 1

Suppose that Player 2 gets a recommendation to play Red. Then,

(a) Player 2 will follow this recommendation blindly.

(b) Player 2 will not follow this recommendation and will play Blue in any case.

(c) Player 2 will ask Player 1 what was recommended to Player 1 and will choose an action depending on the response.

(d) None of the above are true. [*Correct*]

<u>Feedback:</u> "Players cannot communicate. Also they do not follow their recommendations blindly but also do not reject all recommendations. Instead, players investigate whether it is truly worthwhile for them to follow the recommendation using all the

information from the design of the recommendation device. They need information about the recommendation device."

Question 4: Your task in this experiment:

(a) To play four games, several times each, against humans.

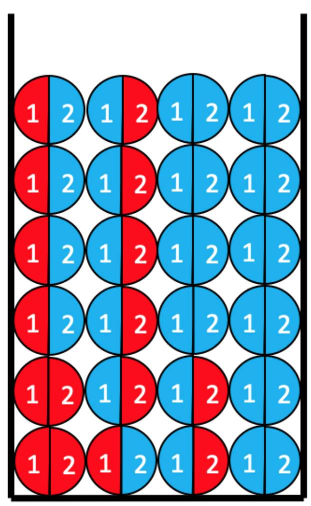
(b) To play four games, several times each, against robots.

(c) To observe and describe how others play the games.

(d) To design recommending devices for given games that will be played by robots. [*Correct*]

Feedback: "As instructions explain, your task is **to design recommendations** to two intelligent robot players on how they should play a game with one another. You will go through a sequence of games and have several chances to improve the device."

Question 5: The following picture displays a recommendation device.



Suppose that one ball is drawn out of this container at random. What is the chance that this ball recommends to player 1 to choose Blue and, at the same time, it recommends to player 2 to choose Red.

(a) 2 out of 24.

(b) 5 out of 24.

(c) 7 out of 24. [*Correct*]

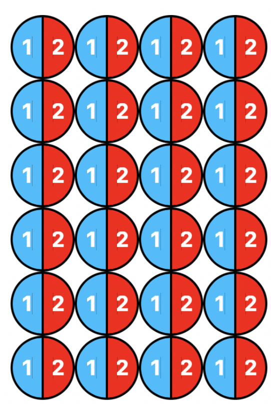
(d) 10 out of 24.

Feedback: "To answer this question, you need to count the number of balls with the left side (1) being Blue and the right side (2) being Red."

Question 6: Suppose that the game that the robot players participate in has the following specific payoffs:

P1 \ P2	Red	Blue
Red	0, 1	3, 3
Blue	2, 5	3, 1

Moreover, suppose that you design the following device:



Which of the following statements is true:

(a) Neither player 1 nor player 2 are willing to follow the recommendation made by the device.

(b) Player 1 and player 2 are both willing to follow the recommendation made by the device and the points you will receive in this round is 0.5.

(c) Player 1 and player 2 are both willing to follow the recommendation made by the device and the points you will receive in this round is 0.5 plus the minimum average payoff received by either robot player equal to 2 points. [*Correct*]

(d) Player 1 and player 2 are both willing to follow the recommendation made by the device and the points you will receive in this round is 0.5 plus the maximum average payoff received by either robot player equal to 5 points.

<u>Feedback:</u> "This device always recommends player 1 to play Blue and player 2 to play Red. With this device, each player will have a certainty about the recommendation to another player. Checking the table of payoffs, we can see that player 1 prefers to follow the recommendation (when player 2 plays Red, it is beneficial to player 1 to play Blue). Similarly, player 2 prefers to follow the recommendation.

According to the instructions, for this round, you would get points based on the smallest of the two average payoffs of players 1 and 2. In this case, player 1 will always (for any of the balls drawn) get 2 and player 2 will always get 5. The smallest of 2 and 5 is 2."

Question 7: Suppose that in one of four games you are presented in the experiment, in the first three rounds player 2 never followed the recommendation of the devices you made.

Which of the following statements is true:

(a) You have a chance to get additional points for this game, only in the case when you design a new device in the fourth round. [*Correct*]

(b) You have no chance to get additional points for this experiment.

(c) You still can earn additional points in the experiment, but not for this game.

(d) If player 2 will follow the recommendation of the device you redesigned for the 4th round in all periods of that round, you will get additional points in that round.

Feedback: "As the instructions explain, for a given game, you receive the number of points you earned in the round for that game with the highest amount of points. This means that if in the last round you will design such device that the recommendations are always followed (by both players 1s and player 2s), you will get points from that round and for this game."

Question 8: Once you design your recommendation device in a round for a given game, the game is played by robot players for multiple periods. Can a robot player choose its action depending on how the game was played in previous periods?

(a) Yes, a robot player observes the actions played in the previous periods and can react to them.

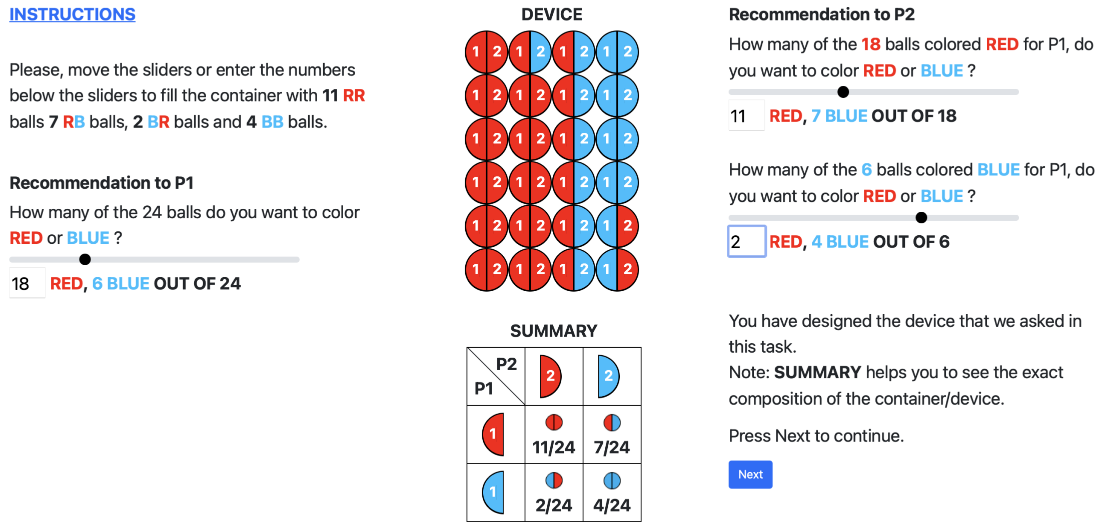
(b) Yes, if it allows a robot player to maximize a total payoff over all periods.

(c) No, in each period, two new robots play the game and they do not have information about the previous periods. [*Correct*]

(d) None of the above is true.

Feedback: "Each game will be played for multiple periods but the robot players in each round will always be different (so that their actions will not depend on what was played in the past). In each period, when a particular color action is recommended to them, they will investigate whether it is truly worthwhile for them to follow this recommendation using all the information from the design of your recommendation device."

Question 9: Please follow the instructions shown on the screen below. [*By moving the sliders, subjects had to implement the distribution* $\mu_{RR} = 11/24$, $\mu_{RB} = 7/24$, $\mu_{BR} = 2/24$ *and* $\mu^*_{BB} = 4/24$.]



Post-Experiment Cognitive Reflection Test (CRT)

- **Question 1:** Together, a mobile phone and laptop cost $1100. The laptop costs 1000 more dollars than the mobile phone. How much does the mobile phone cost in dollars? [*Correct: 50*]

- **Question 2:** It takes 5 bakers 5 hours to bake 5 cakes. How many hours would it take 100 bakers to bake 100 cakes? [*Correct: 5*]

- **Question 3:** In a field there is a group of rabbits. The population of the rabbits doubles in size every day. If it would take 50 days for the rabbit population to completely cover the field, how many days would it take for the rabbits to cover half of the field? [*Correct: 49*]

- **Question 4:** If Charlie drinks one gallon of milk in 3 days, and Emerson drinks one gallon of milk in 6 days, how many days would it take them to drink one gallon of milk together? [*Correct: 2*]

- **Question 5:** Simon decided to invest $2,000 in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. As of October 17, Simon has: [Answers: Lost money in the stock market; Broken even in the stock market; Made money in the stock market] [*Correct: Lost money in the stock market*]

Post-Experiment 11-20 Game Question

Two highly intelligent and competitive individuals are playing the following game. Each player must simultaneously announce an integer number between 11 and 20. Then, the payoffs are determined as follows. Each player receives an amount of dollars equal to the number the player announced. Moreover, if one of the players announces a number that is exactly one less than their opponent, then this player receives an additional reward of $20. Finally, if both players announce the same number, each of them receives an extra $10.

For example, if player 1 says 17 and player 2 says 19, then player 1 receives 17 and player 2 receives 19. Or, if player 1 says 12 and player 2 says 13, then player 1 receives 32 and player 2 receives 13.

What do you think would be the number that one of these players would announce? Or, in other words, if you were one of these two players which number would you announce?

C Feedback Screens

Figs. B1 to B3 display feedback provided to subjects.

D Correlated Equilibria of Tested Games

All 5 games we tested, are presented in the left part of Table 1 in the paper. In this appendix, for each of these games, we derive the set of correlated equilibria, that is distributions over action profiles in $\{Red, Blue\} \times \{Red, Blue\}$ that satisfy Definition 1. We also solve problem (1) and find the desirable correlated equilibrium μ^*. Finally, we show the recommendation device that implements this DCE, i.e., the

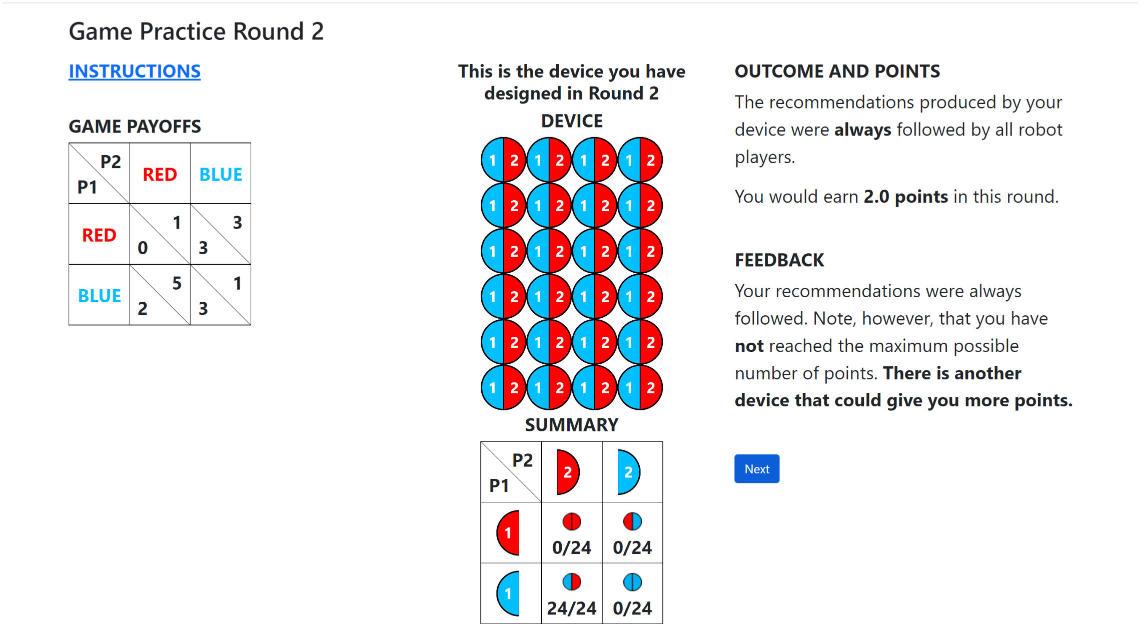


Figure B1: Feedback for a device that implements a sub-optimal correlated equilibrium.

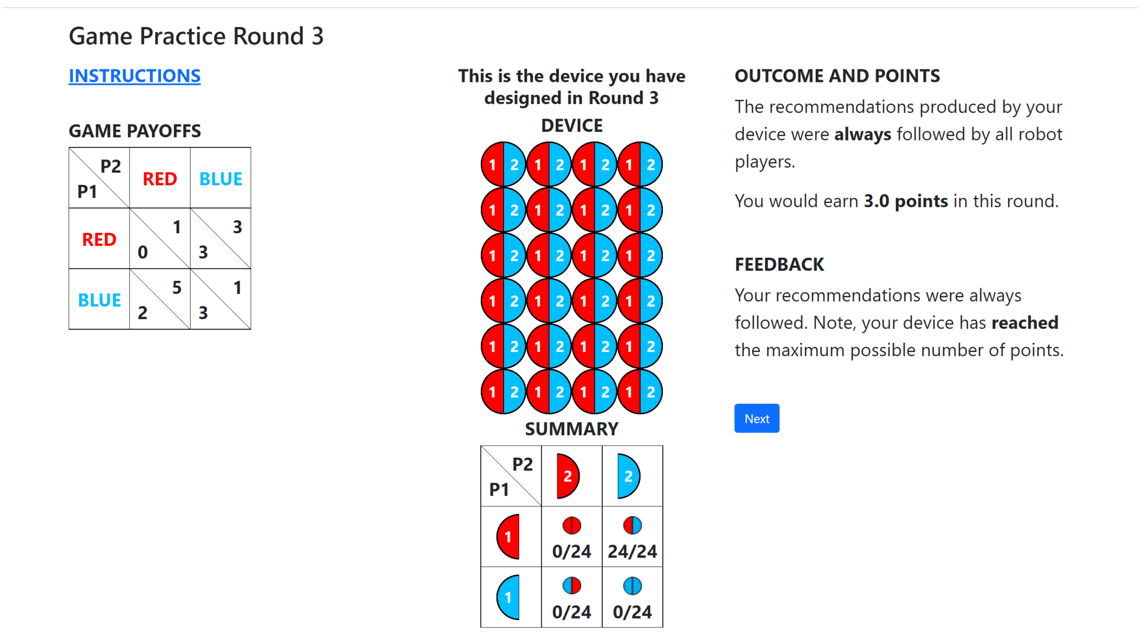


Figure B2: Feedback for a device that implements the desirable correlated equilibrium.

device that gives subjects the maximal payoff in each game.

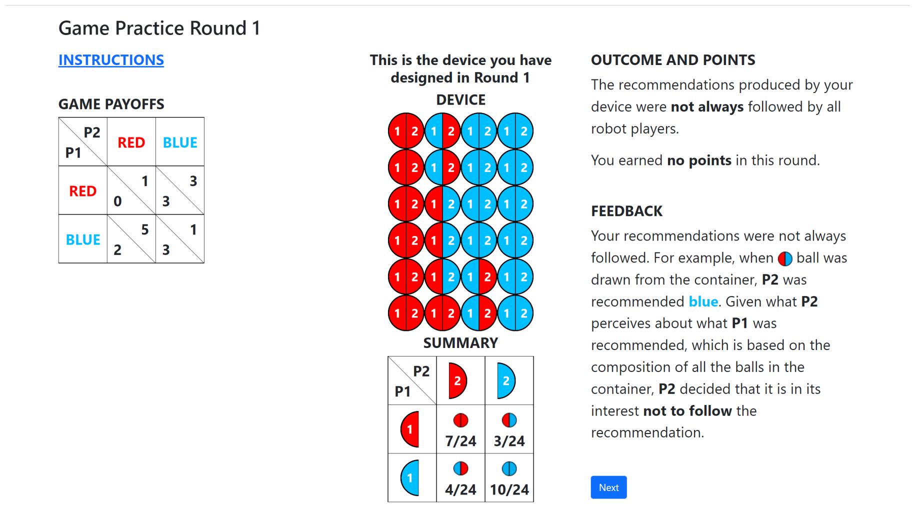


Figure B3: Feedback for a device that is not a correlated equilibrium.

D.1 Prisoner's Dilemma

In Prisoner's Dilemma, each player has a strictly dominant strategy to choose the action Red. As a consequence, in the unique correlated equilibrium (which is, indeed, the unique Nash equilibrium) $\mu_{RR} = 1$ and $\mu_{RB} = \mu_{BR} = \mu_{BB} = 0$. Consequently, this is also the DCE. Thus, the maximum achievable payoff that can be earned in this game is 4. Fig. C1 shows the device that earns this maximal payoff.

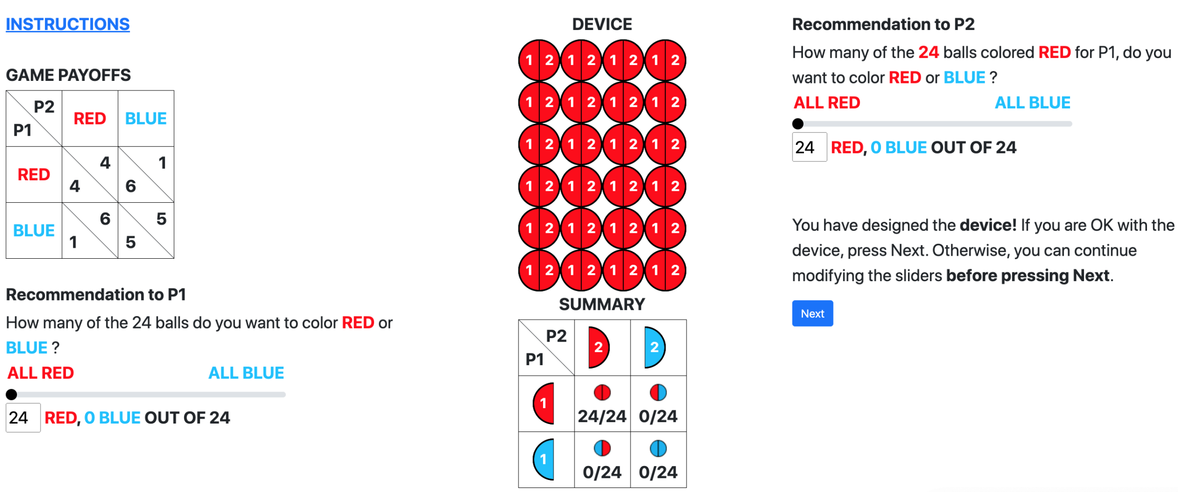


Figure C1: Optimal Device in Prisoner's Dilemma

D.2 Matching Pennies

Using Definition 1, distribution μ is a correlated equilibrium in the Matching Pennies game if it satisfies the following inequalities:

$$5\mu_{RR} + 3\mu_{RB} \geq 3\mu_{RR} + 5\mu_{RB}\,, \qquad 3\mu_{BR} + 5\mu_{BB} \geq 5\mu_{BR} + 3\mu_{BB}\,,$$

$$3\mu_{RR} + 5\mu_{BR} \geq 5\mu_{RR} + 3\mu_{BR}\,, \qquad 5\mu_{RB} + 3\mu_{BB} \geq 3\mu_{RB} + 5\mu_{BB},$$

which can be summarized as

$$\mu_{RR} \geq \mu_{RB} \geq \mu_{BB} \geq \mu_{BR} \geq \mu_{RR}.$$

We conclude that there is a unique correlated equilibrium of this game with $\mu(a) = 1/4$ for all $a \in \{R, B\} \times \{R, B\}$. This is also precisely the unique Nash equilibrium in mixed strategies, where each player independently randomizes over their available actions. Consequently, this is also the DCE and the maximum achievable payoff for this game is $(1/2) \times 5 + (1/2) \times 3 = 4$. Fig. C2 shows the device that would need to be designed in the experiment to earn this maximal payoff.

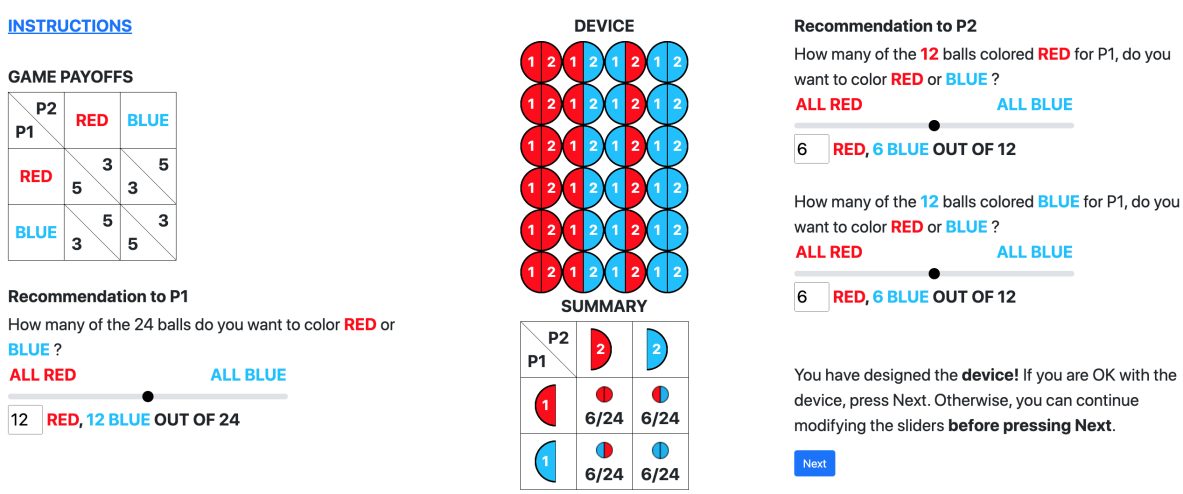


Figure C2: Optimal Device in Matching Pennies

D.3 Battle of the Sexes

Using Definition 1, distribution μ is a correlated equilibrium if it satisfies the following four inequalities:

$$6\mu_{RB} \geq 2\mu_{RR}\,, \quad 2\mu_{BR} \geq 6\mu_{BB}\,, \quad 6\mu_{BR} \geq 2\mu_{RR}\,, \quad 2\mu_{RB} \geq 6\mu_{BB},$$

which is equivalent to the condition that

$$\min\{\mu_{RB}, \mu_{BR}\} \geq \max\left\{3\mu_{BB}, \frac{1}{3}\mu_{RR}\right\}.$$

The DCE solves problem (1). Clearly, at the optimum, $\mu_{BB} = \mu_{RR} = 0$, as the strategies (B, B) and (R, R) result in 0 payoff for both robot players. Finally, to maximize the minimum expected payoff of both robot players, we should set $\mu_{RB} = \mu_{BR} = 1/2$. This achieves a maximum payoff of $(1/2) \times 6 + (1/2) \times 2 = 4$ in this game. Figure C3 displays the device that would need to be designed in the experiment to earn this maximal payoff.

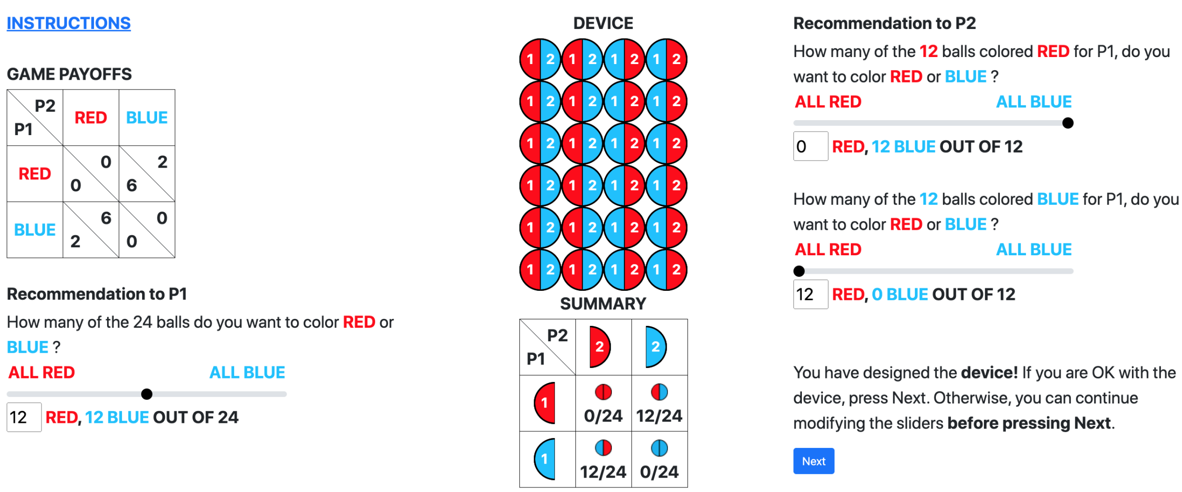


Figure C3: Optimal Device in Battle of the Sexes

D.4 Chicken Version 1

Using Definition 1, distribution μ is a correlated equilibrium if and only if it satisfies the following four inequalities:

$$6\mu_{RB} \geq \mu_{RR} + 5\mu_{RB}\,, \qquad \mu_{BR} + 5\mu_{BB} \geq 6\mu_{BB}\,,$$

$$6\mu_{BR} \geq \mu_{RR} + 5\mu_{BR}\,, \qquad \mu_{RB} + 5\mu_{BB} \geq 6\mu_{BB},$$

which is equivalent to the condition

$$\min\{\mu_{RB}, \mu_{BR}\} \geq \max\{\mu_{RR}, \mu_{BB}\}.$$

Given the payoff matrix, the objective function in (1) is maximized, when as much weight as possible is put on μ_{BB}, as little weight as possible is put on μ_{RR}, and there is a symmetry in the remaining two weights. Thus, the DCE in this case is

$$\mu_{RB} = \mu_{BR} = \mu_{BB} = 1/3,$$

which earns a maximal payoff of $(1/3) \times 1 + (1/3) \times 6 + (1/3) \times 5 = 4$. Fig. C4 displays the recommendation device that yields this maximal payoff.

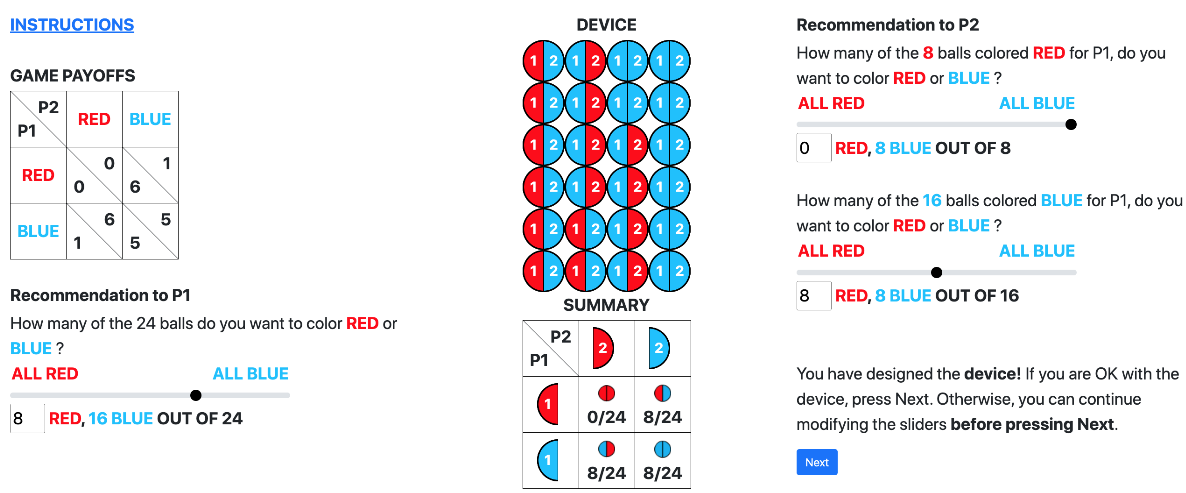


Figure C4: Optimal Device in Version 1 of Chicken

D.5 Chicken Version 2

Using Definition 1, distribution μ constitutes a correlated equilibrium in this case if and only if the following four constraints are satisfied:

$$6\mu_{RB} \geq 2\mu_{RR} + 5\mu_{RB}\,, \qquad 2\mu_{BR} + 5\mu_{BB} \geq 6\mu_{BB}\,,$$
$$6\mu_{BR} \geq 2\mu_{RR} + 5\mu_{BR}\,, \qquad 2\mu_{RB} + 5\mu_{BB} \geq 6\mu_{BB}\,,$$

which is equivalent to the condition

$$\min\{\mu_{RB}, \mu_{BR}\} \geq \left\{2\mu_{RR}, \frac{1}{2}\mu_{BB}\right\}.$$

As in the first Chicken game, the objective function in (1) is maximized, when as much weight as possible is put on μ_{BB}, as little weight as possible is put on μ_{RR}, and there is a symmetry in the remaining two weights. The solution has

$$\mu_{RB} = \mu_{BR} = \frac{1}{2}\mu_{BB},$$

so that $\mu_{RB} = \mu_{BR} = 1/4$ and $\mu_{BB} = 1/2$. This yields a maximal payoff equal to $(1/4) \times 2 + (1/4) \times 6 + (1/2) \times 5 = 4.5$. Fig. C5 displays the device that yields this maximal payoff.

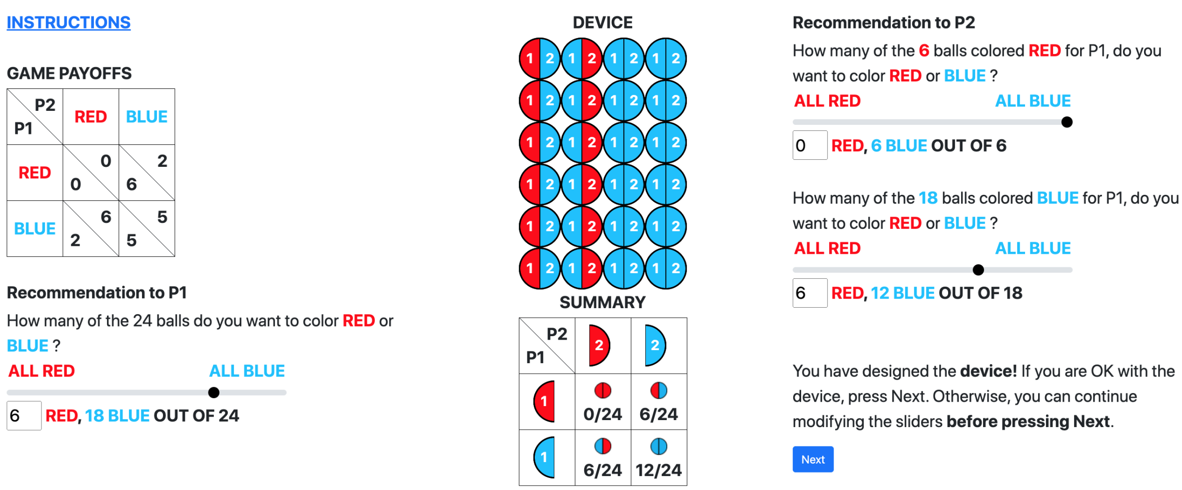


Figure C5: Optimal Device in Version 2 of Chicken