Trick or Treat: Putting Peer Prediction to the Test

Xi Alice Gao and **Andrew Mao** and **Yiling Chen**
School of Engineering and Applied Sciences
Harvard University
{xagao, mao, yiling} @seas.harvard.edu

Abstract

Collecting subjective information from multiple parties is a common problem in collective intelligence. However, incentivizing truthful reports is difficult when there is no ground truth to verify the reports against. *Peer prediction* mechanisms use collected reports alone to provide good theoretical incentives for truthful reporting, but make assumptions that are difficult to satisfy in the real world. They also admit uninformative equilibria where coordinating participants provide no useful information. Using a multiplayer, real-time repeated game, we conduct the first controlled online experiment of a peer prediction method. Our results show that players learn to adopt more profitable strategies through repeated use of the mechanism, and that there is a distinct incentive for participants to converge to the uninformative equilibria.

1 Introduction

Businesses and organizations often face the challenge of gathering accurate and informative evaluation or feedback from separate individuals. A notable example are community-based websites such as Yelp, Amazon, Quora, and Stack Overflow, which are largely dependent on ratings, questions, and answers that are voluntarily contributed by the users of these sites. At the same time, the proliferation of online labor markets has created an opportunity for outsourcing subjective tasks, including classifying images and identifying abusive or adult content on the web, to a readily available online workforce. In all of these settings of collecting subjective information, there exists the significant challenge of incentivizing the participants to honestly contribute their subjective evaluation about some item of interest.

There has been a long line of work on proposing and theoretically analyzing mechanisms to truthfully elicit subjective evaluations from individuals. Yet, in many cases, individual reports cannot be compared to or verified by an observable ground truth. This led to the idea of *peer prediction* mechanisms, where one participant's report is compared to those of his peers to induce a truthful reporting equilibrium among the participants: with proper incentives, it is in a participant's best interest to report his evaluation truthfully if he believes all other participants will also be truthful. Miller, Resnick, and Zeckhauser [7] devised the first such peer prediction mechanism (MRZ mechanism), leading to subsequent mechanisms with good theoretical properties [4, 9, 11, 12].

To the best of our knowledge, there have been few experimental or empirical studies on the performance of peer prediction mechanisms in realistic settings. The properties of these mechanisms reveal several difficulties for doing so. First, the earlier mechanisms often impose strong common knowledge assumptions regarding how participants form beliefs about the item of interest and about other participants' private information. If such a mechanism is used in a setting where these assumptions are not satisfied, then it would be unsurprising if the theoretical guarantees did not hold as well. Moreover, although some peer prediction mechanisms are able to relax these strong assumptions, their payment rules are often described by complicated mathematical formulas. If the participants cannot plainly understand these payment rules, one might wonder if they would actually behave optimally as predicted by the theoretical analyses. Finally, perhaps the most serious problem with peer prediction mechanisms is the existence of uninformative equilibria where participants can blindly agree on their reports without revealing any useful information. This is unavoidable when participants' reports are compared only with each other, and the existence of such equilibria raises questions about the usefulness of these mechanisms; the theory provides little assurance that the participants will choose to play the truthful equilibrium in practice. The above problems inspire two important questions:

- Can the peer prediction mechanisms be adapted and used in a realistic scenario?
- If so, how would participants behave toward peer prediction mechanisms in the presence of multiple equilibria?

We take a first step toward answering these questions by conducting a controlled experiment on the MRZ mechanism via Amazon Mechanical Turk (MTurk), testing it as a repeated game with multiple players. By definition, the MRZ mechanism is a one-shot game. Yet, if peer prediction is used in practice, we might expect participants to interact with and learn about the mechanism through multiple tasks. Hence, we allow the participants to play the game repeatedly so that they may learn to improve their strategies by interacting with others. By having a game with multiplayer, real-time interaction, we capture the learning dynamics of participants when faced with a peer prediction mechanism and

study whether they will converge to one of the multiple pure strategy equilibria of the game.

We make the following contributions:

1. We conduct the first controlled online experiment of the MRZ mechanism through a multiplayer, real-time, and repeated game via MTurk. To our knowledge, this is the first peer prediction method tested in a repeated setting.
2. We formulate a simple story to explain the rules of the peer prediction game to the participants. This story incorporates the required common knowledge assumption in a natural way. We design an intuitive user interface from which the participants can learn to improve their strategies by examining the history of game play.
3. We show that there is a strong incentive for players to converge to the coordinating equilibria, which have higher payoffs than the truthful equilibrium. We present evidence that players can perform simple inference with the prior when choosing their strategies.

Related Work Following the MRZ mechanism [7], several peer prediction mechanisms have been developed to relax the common knowledge assumptions [4, 9, 11, 12], although they often require the participants to make an additional probabilistic report and use much more complicated payment rules to achieve incentive compatibility. Dasgupta and Ghosh [1] study a crowdsourcing setting in which the effort level of a participant determines the probability of observing the correct label, and they propose a mechanism incentivizing both high efforts and truthful reports.

To the best of our knowledge, the only experimental studies on peer prediction mechanisms have been for the Bayesian truth serum (BTS) [10, 8, 3]. The study of Prelec and Seung [8] focus on showing that BTS scores can be used to obtain the ground truth even if most participants' subject judgements are wrong. John, Loewenstein, and Prelec [3] surveyed psychologists about their estimates of questionable research practices and scored them using BTS. This study of the BTS is in a one-shot setting, whereas we test the MRZ mechanism in the setting of a repeated game; we believe that repetition more accurately reveals long-run behavior in practice. Shaw, Horton, and Chen [10] conducted an online experiment using the BTS description as the contextual manipulation for one of the financial incentives tested, but they did not pay the workers according to the mechanism. Gao et al. [2] used the peer prediction method proposed by Witkowski and Parkes [11] to elicit ratings for short tourism ads collected through MTurk, but they simply adopted the method without evaluating it experimentally.

2 Background

We introduce the MRZ mechanism, first proposed by Miller, Resnick, and Zeckhauser [7] and further analyzed by Jurca and Faltings [4].

Consider an item of interest with a finite set Ω of possible types, and let ω be the true type of this item. There are $n \geq 3$ participants who have some experience about this item. The experience of participant i is represented by a private signal s_i drawn from a finite signal space S_i. Each private signal is only observed by the intended participant, not by any other participant or the mechanism. In this work, we assume that the item has a binary type and each participant receives a signal drawn from a common binary signal space.

There is a prior probability distribution $\Pr(\omega)$ over the possible item types ($\sum_{\omega \in \Omega} \Pr(\omega) = 1$). Before the game starts, nature draws the true type of the item according to $\Pr(\omega)$. After that, each participant's private signal is drawn according to the conditional probability distribution $\Pr(s|\omega), s \in S$. A critical assumption of the MRZ mechanism is that the common prior, consisting of both $\Pr(\omega)$ and $\Pr(s|\omega)$, is common knowledge for all participants and for the mechanism.

Once all participants receive their signals, each participant makes a report $r_i \in S$, which may or may not be same as his private signal. Given all participants' reports, the mechanism determines the participants' payments as follows. For participant i, the mechanism randomly chooses one of the other participants as i's reference participant. The reference participant's report is called i's reference report f_i. The payment to participant i, denoted by $u(r_i, f_i)$, is uniquely determined by i's report r_i and i's reference report f_i.

The participants are rational and risk-neutral agents. The payment rule for the MRZ mechanism rule can be derived using strictly proper scoring rules such that truthful reporting is a Bayes-Nash equilibrium (BNE) of the mechanism: a risk neutral participant maximizes his expected payoff by truthfully reporting his private signal if he believes that all other participants will also be truthful. However, for each possible signal $s \in S$, there also exists a *coordinating* BNE in which all participants make the same report s regardless of their private signals. At any such coordinating BNE, the mechanism obtains no information from the participants.

Jurca and Faltings [4] further showed that to sustain the truthful BNE, it suffices for the payment rule to satisfy the following constraints:

$$\sum_{s \in S} \Pr(s|r_i)(u(r_i, s) - u(r_i^{lie}, s)) \geq 0, \forall r_i \in S \quad (1)$$

where r_i^{lie} denotes the signal in S which is not r_i. We use this method to derive the payment rules for our experiment since it allows more choices of parameters for the payment rule. Jurca and Faltings also proved that for the MRZ mechanism in the binary setting, the uninformative coordinating equilibria always exist.

3 The Trick or Treat Story

Perhaps the biggest challenge of studying a peer prediction mechanism in an experimental setting is to present the mechanism in an accessible and intuitive way. This is our main motivation of using the MRZ mechanism: in our setting, the payment rule of the MRZ mechanism consists of 4 parameters and can be presented in a simple table of four rows. Also, each participant need only make a single binary report based on a binary signal, avoiding the potential difficulty of estimating probabilities or continuous values.

The most complicated detail of the MRZ mechanism is the strong common knowledge assumption. To test the

mechanism in a controlled setting, we choose a fixed common prior. Moreover, we create a simple and fun story about trick or treating on Halloween night to incorporate the common knowledge assumption as a natural part of the scenario. We describe this trick or treat story below:

> *A group of kids are trick or treating on Halloween night. There are two types of houses giving out two types of candies, the M&M's and the gummy bears, in different proportions. The kids randomly choose a house to go trick or treating; the house can be one of the two types with equal chance and the kids don't know which type of house was chosen. Each kid secretly and privately gets one randomly selected candy from the chosen house. A clown shows up and asks each kid tell him the type of candy received, promising a payment in return. Each kid may claim to have either type of candy to the clown. To determine the kids' payments, the clown first collects a reports from all the kids. Then for each kid X, the clown randomly selects another kid Y in the group, and kid X's reward is determined by kid X's claim and kid Y's claim according to a table of payment rules.*

The story conveys several important details of the peer prediction game, such as the common prior, the concept of private information, and the concept of misreporting. In our experiment, each player is a 'kid', a particular house is the type of the item, and the candies are the signals and reports. The large supply of candy at each house explains the conditional independence of signals. Finally, the clown is a neutral character that plays the role of the mechanism.

Introduced as part of the background information, the common prior becomes a natural and integral part of the story. This allows us to communicate this potentially difficult concept effectively to the average participants. It was tricky to convey the idea that the proportion of candies given remains the same regardless of how many candies the house has given out (i.e. the signals are conditionally independent) since candies are concrete objects rather than abstract notions. We decided that stating this as a fact was the most straightforward way to convey this idea.

Although the concept of private information and misreporting are common in mechanism design, we were careful to communicate these concepts to MTurk participants in a credible way. Many MTurk workers are wary of the consequences of rejections and blocks, and they typically anticipate what the requester of the task expects and try to submit the correct or the expected answer. Hence, we wanted to ensure that participants would understand that there was no consequence to misreporting their private information. First, we put special emphasis on the fact that each player's private signal is obtained in secret and it is not observed by any other participant or the mechanism. We used the clown as a proxy for the mechanism instead of the requester. In addition, we chose to describe each participant's action as "making a claim", which is a neutral phrasing in lieu of words with negative connotations such as "lying" or "cheating", so as not to invoke the participants' fear of punishments. We also emphasize the fact that each player can claim to have either type of candy, and the clown cannot verify whether the player's claim matches the actual type of the player's candy.

In what follows, we use MM and GB to denote the two possible signals or reports. We refer to the two coordinating equilibria as the MM and GB equilibria.

4 Experiment Setup

To run our experiment online, we use TurkServer [5], a framework and API for conducting online experiments with synchronicity and real-time interaction. In each game, 3 players play repeatedly for 20 rounds. By using a small number of players, we aim to make it easy for each player to reason about other players' actions. At the same time, the large number of rounds gives players sufficient time to explore and improve their strategies.

Each player is paid a $1.00 base payment upon finishing the game. The player also receives a bonus equal to his/her average reward in the 20 rounds of the game (ranging from $0.10 to $1.50). While most tasks on MTurk pay primarily through the base payment, we use the size of the bonus relative to the base payment to motivate workers to pay attention to their reward in the game.

To ensure that no player in a game has prior experience, we limited each worker to participate only once in any experiment. Moreover, we restricted our tasks to US workers, for two reasons. First, our experiment requires synchronicity, and a real-time connection to a US server, so US workers minimize the likelihood of connection issues. Second, controlling for geography avoids unexpected behavior if people from other regions have different behavioral norms or a language barrier in understanding the instructions.

The Task Users progress through the task in several sections. The initial page describes some general information and requires consent. This is followed by an 11-page tutorial, consisting mainly of pictures. The first half of the tutorial describes the trick or treat story, and the second half explains the game interface. After the tutorial, the participant must take a quiz, selecting all correct (true/false) statements out of a total of 14 statements about the task. Each participant has 3 attempts to pass the quiz with a score of at least 80%. If they fail all 3 attempts, they are permanently blocked from our task. After participants pass the quiz, they must wait in a virtual lobby for enough players to start a new game. Whenever there are enough players in the lobby, a 'READY' button appears for each player, and a new game starts when enough players press this button. We explain the game interface in the next section. Finally, participants complete a short exit survey after the game, describing the strategies and reasoning they used.

Interface Figure 1 shows the organization of the main game interface. The top section describes the general information of the game, such as the total number of rounds and the total number of other players. The rest of the interface consists of 2 columns. The left column displays the steps for the current round, and the right column displays the results of the previous rounds and the current round. Other

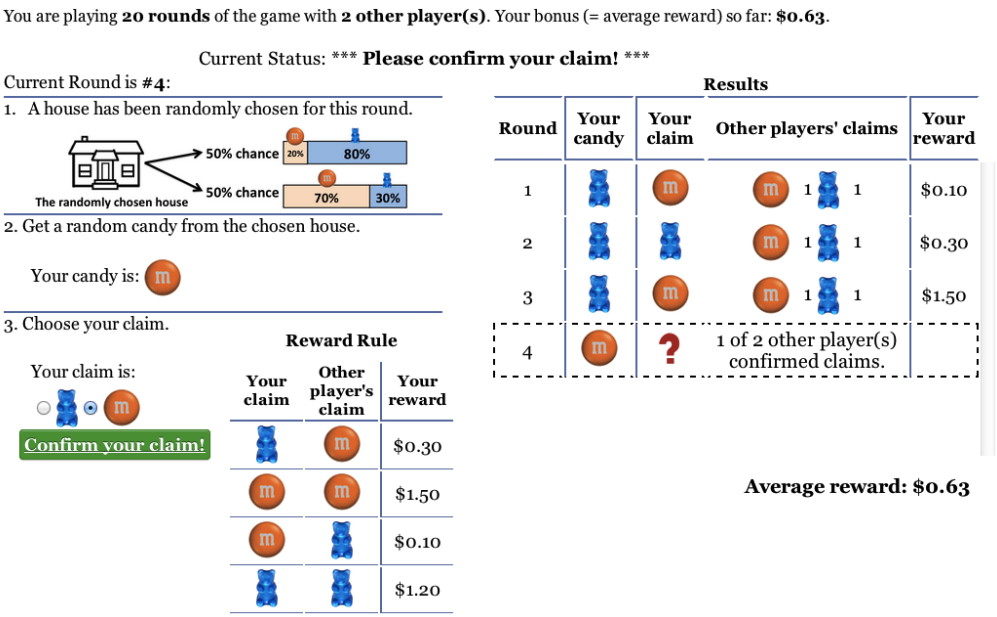


Figure 1: The Game Interface

than viewing previous reports, players cannot communicate with one another during the task.

We carefully considered how to show the right amount of information to participants when designing the right column to show the history of game play. With too much information, a participant may become distracted or confused, and pay less attention to the task. If there is not enough information, a participant may not be able to learn or improve his strategy by observing other participants' actions. After several redesigns, we chose to display the other participants' claims as an aggregate summary for the current participant, since this is the most concise representation that still shows all other participants' actions.

To control for position biases on the game interface, we randomize the row order of the payment rule once for each participant and show this randomized table throughout the task. We also randomize the order of questions on the quiz and order of the radio buttons for choosing claims.

Dealing with errors In a typical MTurk task, workers expect the task to progress smoothly as long as they are attentive. However, in a synchronous experiment players can disconnect or experience other technical issues. To ensure that a game progresses smoothly when such issues occur, we expel a participant from the game if disconnected for at least 1 minute (a reasonable threshold since a typical game takes less than 5 minutes to finish). An expelled player cannot reconnect to the game, and the server will choose truthful reports on behalf of the expelled player. This ensures that other players experience the game as normal, but we exclude all games with expelled players in our analysis because of the fake replacement.

Treatments We conducted two experimental treatments using different payment rules. In the first treatment, we would like to answer the following question:

- Given a typical payment rule, will the players learn to play one of the pure strategy equilibria and if so, which equilibrium will they converge to?

The prior and the payment rule used are shown in Equation 2 and Table 1 respectively. We construct the payment rule such that all four combinations of report and reference report result in different payoffs. Also, a report of MM may result in either the maximum or the minimum payment, so neither report dominates the other in terms of the possible realized payments. In terms of maximizing payoffs, the MM equilibrium seems to be the most favorable choice. However, under this prior, if one player receives a given signal, it is more likely that another player also receives the same signal. This simple reasoning may influence players to be truthful if they believe other players are also truthful.

$$\Pr(A) = 0.5, \Pr(MM \mid A) = 0.2, \Pr(MM \mid B) = 0.7 \quad (2)$$

	ref report = GB	ref report = MM
report = GB	$1.20	$0.30
report = MM	$0.10	$1.50

Table 1: Typical Payment Rule

For the second treatment, we would like to test whether making the payoffs for the two coordinating equilibria equal can deter the players from reaching either coordinating equilibrium, especially when they cannot communicate with one another. We use the same prior as the first treatment and change to a simpler payment rule rewarding agreement between the two reports, as shown in Table 2. A participant receives the maximum payment if his report agrees with the reference report, and the minimum payment otherwise.

5 Results and Challenges

We collected the results of 103 and 104 games (excluding games with expelled players) for the two treatments re-

	ref report = GB	ref report = MM
report = GB	\$1.50	\$0.10
report = MM	\$0.10	\$1.50

Table 2: Payment Rule Rewarding Agreement

spectively. We characterize the games converging to each pure strategy equilibrium, and propose a statistical model of strategies to analyze the learning effect throughout the game.

We received generally positive feedback about the design of our task and its difficulty, clarity, and enjoyability. While we were initially concerned about the complexity of the task, 81% of workers who attempted the quiz eventually passed it before being locked out. This suggests that the quiz was of appropriate difficulty and that most workers were able to understand the tutorial. Moreover, participants provided positive feedback about the task in their exit survey comments, claiming that the game was easy to understand, quick, smooth and enjoyable. These observations suggest that, with careful design of the interface, the MRZ mechanism can indeed be made accessible to most people.

Equilibrium Convergence The usefulness of peer prediction in practice depends on a simple question: *which one of the multiple pure strategy equilibria (i.e. truthful, MM, and GB) will the players converge to and why?*

We first examine this convergence with a simple method. Let t be a particular strategy under consideration (truthful, MM, or GB.) To determine if all players in a game converged to playing strategy t reasonably early, we find the earliest round d_i^t such that player i's actions from round d_i to round 20 are all consistent with strategy t, and we take round $\max_i d_i^t$ to be the round at which all players converged to using strategy t. Note that due to the realization of signals, a sequence of actions may be consistent with more than one pure strategy. Finally, we compute $t^* = \arg\min_t(\max_i d_i^t)$. We chose 15 to be the threshold for determining whether the convergence occurred early enough in the game, If $t^* \leq 15$, then we consider the game to have converged to the pure strategy equilibrium with strategy t^*. Otherwise, the game remains unclassified. We use this method to classify the games, and the results are shown in Table 3.

	MM	GB	Truthful	Unclassified	Total
Treatment 1	47	4	5	47	103
Treatment 2	7	34	7	56	104

Table 3: Simple equilibrium convergence classification

The simple method is not robust if players deviate and explore other strategies. To account for these explorations, we use a relaxed method that allows for up to 3 reports to deviate from the particular strategy when determining round d_i^t (3 is small enough to limit the extent of allowable deviations while allowing more games to be classified). The new cutoff threshold is 15 minus the number of deviated reports (up to 3). This relaxed method allows us to classify more than 75% of games in each treatment, as shown in Table 4.

Tables 3 and 4 show similar results. For treatment 1, a majority of the games converged to the MM equilibrium, as expected. Yet, to our surprise in treatment 2, the majority of the

	MM	GB	Truthful	Unclassified	Total
Treatment 1	62	11	18	12	103
Treatment 2	12	47	22	23	104

Table 4: Relaxed equilibrium convergence classification

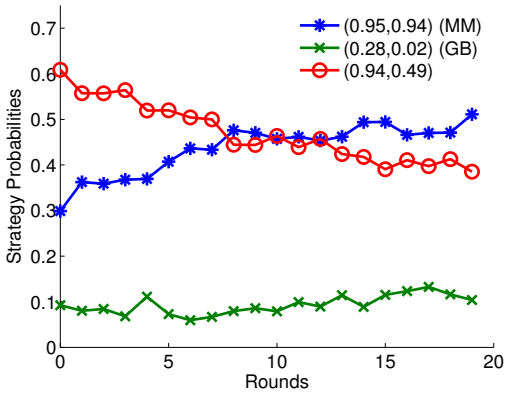
games converged to the GB equilibrium instead of the truthful equilibrium. The total number of games converging to a coordinating equilibrium in treatment 2 is comparable to that of treatment 1, suggesting that giving the two coordinating equilibria equal payoffs did not appear to significantly deter the participants from choosing them. Participants' comments revealed that they chose the GB equilibrium in treatment 2 because the probability of receiving the GB signal is greater than that of the MM signal. Together, these two treatments point toward a strong incentive to choose the coordinating, uninformative equilibria when they result in higher payoffs than the truthful equilibrium.

Learning through rounds We use a probabilistic model of strategies to further investigate the learning effect throughout the repeated game. First, using all available data for each treatment, we use the expectation-maximization (EM) algorithm to simultaneously estimate a set of K mixed strategies and a prior probability distribution over these strategies. For each round of each game, we assume that a mixed strategy is drawn based on the prior distribution, and that the three players' reports are generated independently from this strategy. Then, for each round, we estimate the posterior probability distribution of this set of strategies using all games in that round. This allows us to observe how the distribution over these strategies changes over successive rounds.

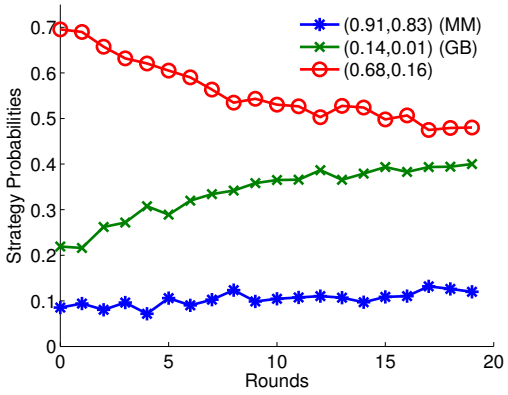
Given the three pure strategy equilibria, we choose $K = 3$ and plot the posterior distribution over the strategies in Figures 2a and 2b. A strategy in the figure is denoted by $(\Pr[r_i = MM \mid s_i = MM], \Pr[r_i = MM \mid s_i = GB])$, the probability of reporting MM given the MM and GB signals respectively.

Each strategy that we learned can be associated with a pure equilibrium strategy. The MM and GB strategies are quite apparent, while the third strategy in red is closest to truthful. In both figures, the coordinating strategies show upward trends, with one being clearly dominant (the MM strategy in Figure 2a and the GB strategy in Figure 2b). This shows the same conclusion as our previous analysis, and in both cases the truthful strategy is crowded out. Hence, as the game progresses, more players are adopting the MM or the GB strategies, which have higher payoffs than the truthful equilibrium.

Experimental Challenges To carry out our experiment on MTurk, we used methods building on those used by Mason and Suri [6]. On MTurk, a requester typically posts some tasks and waits for workers to complete these tasks. This does not work well for our experiment, for several reasons. Since each game requires multiple workers to participate simultaneously, workers will need to wait in the lobby for a long time if they accept our tasks at very different times. If this happens, games may start with players not paying attention, causing further frustrations for other players. The re-



(a) Aggregate strategies in treatment 1.



(b) Aggregate strategies in treatment 2.

Figure 2: Estimated strategies & distribution over rounds.

quirement of unique workers further exacerbates this problem. Since each worker participates only once, there is no way to contact them beforehand (for example, to recruit a group for a specified time).

To solve this problem, we used an idea similar to the recruitment process for lab experiments. Typically, subjects sign up for a lab experiment with their email addresses and will be notified of the time and place for experiments. Similarly, we create a separate recruitment task where a worker consents to participation and provides specific times of day when he/she is available to participate. Once many workers have completed this recruitment task, we schedule experiments at ideal times and invite them to participate through email. For each specified time, we post tasks during a limited time window, to encourage the timely arrival of the participants and avoid long waiting periods.

This recruitment process worked extremely well for our experiment, collecting data for the required number of games significantly more quickly than running ad hoc experiments. We also observed many fewer games with connection issues, giving better data quality as a result.

6 Discussion

Our experiments are the first online experiment of the MRZ peer prediction mechanism through a multi-player, real-time, repeated game. We demonstrate that the MRZ mechanism is accessible to laypeople by explaining its details in a simple and fun story. Using an intuitive user interface, we studied how the mechanism affects players' strategies over repeated rounds, and discover that there is a strong trend for players to choose the coordinating equilibria, with higher payoffs, over the course of the game. This raises questions about the usefulness of peer prediction mechanisms in realistic settings where users may participate multiple times.

Given the strong incentives to choose the uninformative equilibria, a clear future direction is to explore different techniques, ranging from social or psychological to technical, for influencing players to choose the truthful equilibrium. For example, the interface could be augmented to display statements about the truthful equilibrium, or we could add honest artificial players to each game. Our results also suggest that participants can perform simple inference using prior information, but how much does the prior actually affect their decision? We can imagine testing this by examining behavior when the payment rule does not induce a truthful equilibrium for the given prior.

References

[1] Dasgupta, A., and Ghosh, A. 2013. Crowdsourced judgement elicitation with endogenous proficiency. *ACM International World Wide Web Conference (WWW)*.

[2] Gao, X. A.; Bachrach, Y.; Key, P.; and Graepel, T. 2012. Quality expectation-variance tradeoffs in crowdsourcing contests. In *AAAI*.

[3] John, L. K.; Loewenstein, G.; and Prelec, D. 2012. Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science* 23(5):524–532.

[4] Jurca, R., and Faltings, B. 2009. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research* 34(1):209–253.

[5] Mao, A.; Chen, Y.; Gajos, K. Z.; Parkes, D.; Procaccia, A. D.; and Zhang, H. 2012. TurkServer: Enabling Synchronous and Longitudinal Online Experiments. In *Proceedings of the 4th Workshop on Human Computation (HCOMP'12)*.

[6] Mason, W., and Suri, S. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods* 44(1):1–23.

[7] Miller, N.; Resnick, P.; and Zeckhauser, R. 2009. Eliciting informative feedback: The peer-prediction method. *Computing with Social Trust* 51(9):1359–1373.

[8] Prelec, D., and Seung, H. An algorithm that finds truth even if most people are wrong.

[9] Prelec, D. 2004. A Bayesian truth serum for subjective data. *Science* 306(5695):462–466.

[10] Shaw, A.; Horton, J.; and Chen, D. 2011. Designing incentives for inexpert human raters. *Proceedings of the ACM 2011 conference on Computer supported cooperative work* 275–284.

[11] Witkowski, J., and Parkes, D. C. 2012. Peer Prediction without a Common Prior. *EC 2012* 1–18.

[12] Witkowski, J. 2012. A Robust Bayesian Truth Serum for Small Populations. *AAAI 2012*.