What Data do Players Rely on in Social Deduction Games?

Sarah Wiseman
Goldsmiths, University of London
London, UK
s.wiseman@gold.ac.uk

Kevin Lewis
Goldsmiths, University of London
London, UK
klewi014@gold.ac.uk

Abstract
The prisoner's dilemma is a well-studied game, with much information available about successful strategies. In this work we explore how players make decisions within an iterative version of the game, played against varying opponents in a closed environment. We present an app-facilitated social deduction game for eight or more players based upon the prisoner's dilemma (PD). Players sat along a table and played PD, with the results causing players to move up or down the table, potentially facing a new opponent each round. We highlight which sources of information were most important to players when deciding whether to betray or cooperate, as well as deciding whether their opponent was trustworthy. Based upon player-reported data, we find players used five sources of information in the social deduction game. Data suggests that facts provided to players about their opponents are often overlooked in favour of information gained during social interaction or from events that occurred previously in the game.

CHI PLAY EA '19, October 22–25, 2019, Barcelona, Spain

ACM ISBN 978-1-4503-6871-1/19/10.
https://doi.org/10.1145/3341215.3356272

Author Keywords
Social deduction game; trust game; game theory; strategy; prisoners dilemma; decision making.

CSS Concepts
• Human-centered computing~Empirical studies in collaborative and social computing

Introduction
Social deduction games are a popular genre of table-top games. Players aim to determine the correct course of action based on potentially misleading data given to them by their fellow players. Knowing when to trust an opponent and when to be wary is the key to success. But how do players make this choice? What sources of data do people rely on when making such decisions in social deduction games? In this work, we explore how

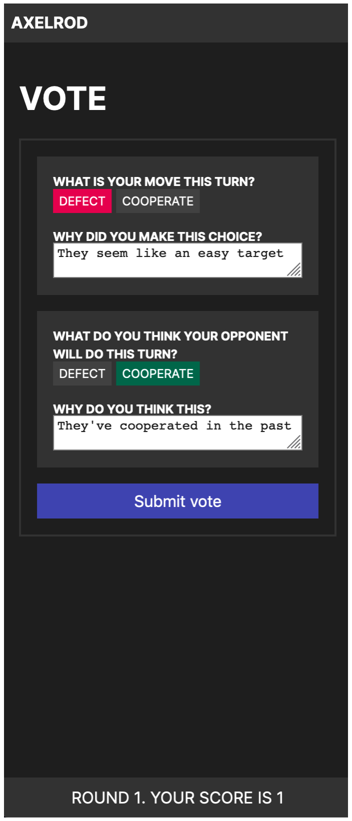

Figure 1: Axelrod player interface. Players are asked to choose a move and justify it and guess what their opponent will do and justify it. The round counter and score are shown at the bottom of the screen.

players make decisions in a social deduction game and which sources of information in the game help players to make those choices.

To investigate this question, we developed the game "Axelrod", a purpose built, app-facilitated social deception game based upon the Prisoner's Dilemma. In the game, players sit opposite their opponent in two long lines either side of a table. Each round, the players have a discussion with the player across the table and then must choose to cooperate or defect. Based on the results, players gain or lose a number of points and potentially move up or down the table based on their place in the leader board. Players decisions throughout the game are collected through a web app running on each players' phone. The app collected data about cooperation/defection, player scores, and prompts the players to report on why they have made each decision. Each side of the table may move in an asymmetrical fashion, resulting in new player pairings for the following round. The winner is the player in the top seat with the most points at the end of the final round.

At the half-way point, some players are shown selective facts *(Game facts)* about the game, such as their opponent's voting history, their opponent's justifications for decisions or the score of the person one seat up or down from them. This added to the potential data sources that players had at their disposal when deciding on the strategy they would use during the game.

During each round, participants have a number of choices available to them. Do they want to betray or cooperate with their opponent? Do they want to attempt to deceive their opponent? Do they trust what their opponent is telling them? Past research has shown that whilst personality can determine a player's decisions within the game [4], a number of external social cues can also affect these decisions including having empathy for your opponent [3] or fearing ostracism from a group [9].

In this paper we aim to explore which information players choose to rely on most when playing a social deduction game – will players rely upon the facts provided by the app, or will they rely upon their own intuition and social interaction?

Related Literature

The Prisoner's Dilemma

The Prisoner's Dilemma (PD) is a game which is often used to model a strategic interaction between two players, and demonstrates why two players may not cooperate even when doing so looks like the most rational choice. In the game, players can choose to either cooperate or defect. A defection results in a higher payoff than cooperation, unless both players defect in which case both do worse than if they had cooperated. When played iteratively, the players can start to learn their opponents' strategies and can begin to inform their own decisions based on this information (See [2] for a review of strategies).

When a game of iterated prisoner's dilemma is played by human agents, there are additional factors that may impact the decision to cooperate or defect. Research has shown that characteristics such as empathy can increase levels of cooperation in game [3].

Social Deduction in Entertainment

The Betrayer's Banquet [1] is an experiential dining game based on the iterated prisoner's dilemma. Players are invited to sit at a dining table for a meal. The food at the head of the table is significantly nicer than that offered at the foot of the table. Diners play a round of the Prisoner's Dilemma with the player across the table and then move up or down depending on whether they cooperated or betrayed. This experience highlights how the iterative Prisoners Dilemma can be used to create an entertaining game within a group.

In The Justice Syndicate [6] - an interactive performance piece by ex-theatre company fanSHEN – players are based in a "courtroom" where they must act as jurors, deciding the fate of an accused surgeon. Players make decisions based upon the information given to them via an app which supports the game. Game admins run real-time analysis to determine which information should be provided to certain players, in an attempt to alter their verdicts. This work suggests that player choice can be altered by being given facts about the situation in which they find themselves a player.

Decision making in Social Deduction Games

The Ultimatum game, like PD is an example of a social cooperation game. Research has suggested that having information about an opponent's past actions can affect a player's decisions [5,10], as can being given general information on common strategies [11]. Other research has shown that limiting the topics of communication for players in the pre-game phase can lead to different decision making [12]. This work suggests it is possible to alter the players' decision-making process by changing certain conditions.

Method

A web-based application was developed to support the game (See figure 1). The app had a number of features within the game:

1. Present players with the choice to cooperate or defect
2. Ask players why they made that decision
3. Ask players to predict what their opponent would do and provide reasoning
4. Show game facts to players

Procedure

Players were randomly allocated a seat around a long table. There were equal numbers of players on each side of the table, and each player had a clearly assigned "opponent" seated opposite them.

Players were asked to navigate to the game app via a browser on their phone. In the first instance, players were asked to choose a username for themselves. The game master then explained the rules of the game to players and then the rounds of the game began.

Players were allowed 1 minute to have a discussion with their opponent in each round. They were encouraged to use this time to "work out" what their actions they would take. Players were then asked to choose to cooperate or defect. To indicate this, users would input their choice on the app. This step was done in silence and took up to 30 seconds to complete.

Half way through the games, half of the players were given a fact about their new opponent at the start of the round via the app. The type of fact provided was randomly chosen. There were six fact types offered:

- *Last 3 rounds* told the player what their opponent had done in the last 3 rounds and their reasoning.
- *Full voting history* showed the opponent's voting history for the entirety of the game.
- *Full seat history* showed the opponent's seat positions throughout the game.
- *Percentage* showed an overall percentage of how often the opponent had defected.
- *Score of neighbouring seats* showed the scores of players either side
- *Opponent score* showed the opponent's current score.

Players were given scores varying from -2 to 2 depending on the result of their actions. Figure 2 shows these allocations. Depending on their overall score, players were moved up and down the table so that all players were sitting in score order. Players did not move across the table.

Participants
Four games of Axelrod were played, Game 1 was eight rounds long, after which it was decided that more rounds could be added to make the game more engaging. Games 2 and 3 were 14 rounds long. Game 4 was eight rounds long due to time constraints.

Games 1, 2 and 3 had 10 players, Game 4 had 14. A total of 33 distinct players were involved in the games (11 players took part in more than one game). Players ages ranged from 18 to 45+. Some players had experience playing social deception games previously (eg Werewolf) whereas others had none.

Results
The percentage of cooperate decisions across all four games ranged from 65% to 84% (*M=71%, SD=10%*). Players in the game expected their opponents to vote to cooperate between 57% to 66% of the time (*M=61%, SD=4%*).

Player A \ Player B	**Player B Defects**	**Player B Cooperates**
Player A Defects	Both players -1 point	Player A +2 points Player B -2 points
Player A Cooperates	Player A -2 points Player B +2 points	Both players +1 point

Figure 2: Defection matrix for players in Axelrod

Players guessed their opponent's move correctly between 66% and 81% of the time (*M=72%, SD=4%*). In cases where a player had incorrectly predicted their opponent's move, in 48%-86% (*M=68%, SD=17%*) of cases, the player guessed their opponent would defect when in fact they cooperated.

Evaluation
Strategies
Players appear to be pessimistic in their expectations of their opponents – more frequently when a player predicted their opponent's actions incorrectly, it was because they expected their opponent would defect when in fact they cooperated.

It was difficult to determine strategies that players had throughout the game. Some players reported a tit-for-tat style of play stating that they would "Cooperate until betrayed" (*Game 3 Player E*). Many players appeared to use changing strategies depending upon their location on the table.

Despite being given no further advice than that the top of the table was the "best" and the bottom the "worst", some players appeared to take it upon themselves to form short, dynamic teams with their opponent. Multiple players reported in game that they were either both trying to move up the table together, or players were making strategic decisions for one to defect whilst the other cooperated in order to increase the point delta with a player in a seat below. One player stated their reason to betray their opponent was an "agreement to screw over the guy to the left" (*G4PF*), a strategy which was confirmed by their opponent.

This team-forming behaviour led to a new strategy that we observed in this game which was a result of the iterative, yet multiple opponent ecosystem that was used in the game. The "Protect" strategy is used by a player when they are sat at the top of the table and want to protect their seat. This resulted in players cooperating with one another until another player in the seat below got too close in points, at which point the two top players would adapt a betray strategy to get one of them the points needed to get away from the contender. This led to players making decisions in order to "protect [their] position" *(G3PA)*.

At the other end of the table there were player "teams" who had decided that neither of them were getting out of the bottom seat, and as such it would be fun to try and get as few points as possible, thus taking the "Always Defect" strategy in order to minimize their scores. With G4PJ explaining they were taking part in "a suicide pact basically".

Making Decisions

Players were asked to report on how they had made the decision to defect or cooperate after the votes had been cast but before they had been revealed. These short text answers were coded by two researchers separately to understand what information players were using when making decisions in the game.

After coding the data, five sources of information emerged from the player reported data: Current data, Observed Data, Historical Data, Meta Data and Provided Data. Below we elaborate on each data source and provide a figure explaining how many decisions in the 4 games relied upon this particular data source.

Current Data (23.3%) was data that players has gleaned from the conversation with their opponent prior to voting that round. This would include simple reasons such as their opponent telling them what their choice would be, "He claimed that he will" *(G1PB)* to sharing strategies "Neither of us want negative points" *(G3PC)*.

Historical data (8.1%) was used when a player made a decision based on previous events in the game. One example of this was when a player had faced an opponent previously in the game and remembered their actions. This worked both ways, some players decided to seek revenge for previous actions "She has already betrayed me" *(G3PE)* whereas others chose to trust a player based on a previous positive interaction "We cooperated last time" *(G4PT)*. Other players used historical data on their own performance to help them make the decision "I keep losing" *(G1PA)*.

Meta data (4.4%) included information that players had obtained from outside the game itself. This might include information from previous games, or social interactions between players who knew each other outside of the game. In these instances players could make assessments about a person's trustworthiness based on their prior social interactions "Who would trust [player]?" *(G1PL)*. It was not always the case that players allowed social history to affect the game play, some were keen to not let game play affect their relationship "we chose to do it out of trust for our friendship's history" *(G2PS)*.

Observed Data (3.4%) included information that players had determined based on the movement of their opponent during the game. If suddenly they observed their opponent move up multiple seats to meet them, then they could determine that it was likely the opponent had just betrayed the previous player they were set against "I know he [betrayed] last time-he went up 2" *(G4PN)*.

Provided data (2.3%) were the facts given to players via the app during the game (see Method). For example, a player was told via the app about their opponent's strategy of always defecting and decided they would have to take a similar approach "Because they are defecting and I'm taking them down" *(G3PL)*.

Primarily players appear to rely on *Current data* to decide how to play that round, using the conversations that they have just had with the player to determine whether or not they could be trusted. *Historical data* were used less frequently, but players still clearly relied on this in-game information to inform their decisions. *Meta data* and *Observed data* were infrequently used, and was not something that could be controlled within the game by the game master (GM). The data set that could be controlled by the GM was the *Provided Data*. This data was also not commonly used by players, with very few of them reporting that the facts they were shown affected their decisions.

Conclusion

Whereas we hypothesized that providing objective data to players might affect the strategies they chose, it appears more commonly that players relied upon their own experiences in the played game.

The cause of this might be that players did not trust the facts that were provided to them, preferring to use their own data sources. Or it could be that the facts provided to them were not used consistently (they were only given to half of players for half of the game) meaning that players did not get used to checking them.

Future Work

Future work will aim to validate these data sources in other technologically mediated social deduction games to explore whether they are seen in games beyond those set around the Prisoners Dilemma.

We also aim to explore how the *Provided Data* can be presented differently to encourage more user engagement. And subsequently to understand why users choose to prioritise some sources of data over others.

References

1. Jacob Aron. 2013. How the prisoner's dilemma changes diners' etiquette. *New Scientist*. Retrieved from https://www.newscientist.com/article/dn24276-how-the-prisoners-dilemma-changes-diners-etiquette/.

2. Robert Axelrod. 2009. Effective Choice in the Prisoner' s Dilemma. *The Journal of Conflict Resolution* 24, 1: 3–25.

3. C Daniel Batson and Nadia Ahmed. 2001. Empathy-induced altruism in a prisoner's dilemma II: what if the target of empathy has defected? *European Journal of Social Psychology* 31: 25–36.

4. Christophe Boone, Bert De Brabander, and Arjen Van Witteloostuijn. 1999. The impact of personality on behavior in five Prisoner's Dilemma games. *Journal of Economic Psychology* 20, 3: 343–377.

5. Juergen Bracht and Nick Feltovich. 2009. Whatever you say, your reputation precedes you: Observation and cheap talk in the trust game. *Journal of Public Economics* 93, 9–10: 1036–1044.

6. Rachel Briscoe. 2017. Near Now - The Justice Syndicate. *Near Now*. Retrieved February 2, 2019 from https://nearnow.org.uk/stories/the-justice-syndicate.

7. Gokul Chittaranjan and Hayley Hung. 2010. Are you a Werewolf? Detecting deceptive roles and outcomes in a conversational role-playing game. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*: 5334–5337.

8. Markus Eger and Chris Martens. 2018. Keeping the Story Straight: A Comparison of Commitment Strategies for a Social Deduction Game. AIIDE: 24–30.

9. David Hirshleifer and Eric Rasmusen. 1989. Cooperation in a Repeated Prisoners ' Dilemma with ostracism. *Journal of Economic Behavior and Organization* 12, 1: 87–106.

10. Martin A. Nowak, Karen M. Page, and Karl Sigmund. 2000. Fairness Versus Reason in the Ultimatum Game. *Science* 289, 5485: 1773–1775.

11. Alan G. Sanfey. 2009. Expectations and social decision-making: Biasing effects of prior knowledge on Ultimatum responses. *Mind and Society* 8, 1: 93–107.

12. Ro'i Zultan. 2012. Strategic and social pre-play communication in the ultimatum game. *Journal of Economic Psychology* 33, 3: 425–434.