Australian
National
University

Higher-order Beliefs in a Sequential Social Dilemma

by

Evan M. Calford

and

Anujit Chakraborty

ANU Working Papers in Economics and Econometrics
681

December 2021

JEL Codes: C92, D81, D91

ISBN: 0 86831 681 4

HIGHER-ORDER BELIEFS IN A SEQUENTIAL SOCIAL DILEMMA

EVAN M. CALFORD* AND ANUJIT CHAKRABORTY**

ABSTRACT. Do experimental subjects have consistent first and higher-order beliefs about other's preferences? How does any inconsistency affect strategic decisions? We introduce a simple four-player sequential social dilemma where actions reveal first and higher-order beliefs. The unique sub-game perfect Nash equilibrium (SPNE) is observed less than 5% of the time, even though our diagnostic treatments show that a majority of our subjects are self-interested, higher-order rational and have accurate first-order beliefs. In our data, strategic play vastly deviates from Nash predictions because first-order and higher-order beliefs are inconsistent for most subjects.

Keywords: Experimental economics, Higher-order beliefs, Social dilemma.
JEL codes: C92, D81, D91

* Research School of Economics, Australian National University; evan.calford@anu.edu.au.
** Department of Economics, University of California, Davis; chakraborty@ucdavis.edu.
We thank seminar audiences at UCL, UC Davis, Purdue University, IIMA, Academia Sinica, U of Arizona, WZB, the ESA World Meetings and the East Anglia Behavioral Game Theory Workshop. Conversations with Rod Garratt, Burkhard Schipper, Amanda Friedenberg and Ala Avoyan are gratefully acknowledged, as is detailed feedback from Alex Brown. We thank Phillip Chapkovski for providing advice and sample code for software implementation. Chen Wei and William Brown provided very helpful research assistance. Funding for subject payments was provided by the John Mitchell Economics of Poverty Lab, Australian National University.

> "If you know the enemy and know yourself you need not fear the results of a hundred battles." ∼ Sun Tzu

Everyday strategic decisions are inherently risky: we know little about how our opponents would act. Assuming that opponents act rationally subject to their preferences only pushes the uncertainty one level up: we know little about their preferences. Knowing the "enemy", as recommended by Sun Tzu, is unfortunately no easy task! A game theoretic analysis under incomplete information typically resolves this issue by assuming that all players agree about the probability distribution over other's cardinal preferences (i.e. utility). But, for experiments run in the laboratory or in the field there is little reason to believe that strangers, who have had different life-experiences prior to that interaction and who have never communicated with each other, would hold identical distributions over other's preferences. This raises the following questions: Are typical experimental subjects themselves aware that they might mutually disagree about the distribution of preferences? If they are aware, then how does their understanding of this disagreement influence their strategic behavior?

We refer to what i believes about the distribution of other's preferences as i's first-order beliefs, and refer to what i believes j believes about the distribution of other's preferences as i's second-order beliefs. The collection of all such multiple iterations (e.g., what i believes j believes $k \neq j$ believes...) are referred to as i's *higher-order beliefs*, in contrast to i's *first-order beliefs*. We introduce a four-player sequential game with *perfect information*, called the Sequential Social Dilemma game (SSD henceforth), that independently identifies these first- and higher-order beliefs through revealed preference.

In the SSD game, four players, P1 to P4, each of whom occupy a red node, sequentially choose between a *safe action* and one of three *risky actions*. The safe action, which we present to subjects as forming a link to a blue node, earns a fixed return of 24. A risky action, presented to subjects as linking to one of the three other red nodes, earns a payoff that depends on the linked player's action. A risky action pays 30 or 10 if the linked player plays the safe or a risky action, respectively. In the Subgame Perfect Nash Equilibrium (SPNE) with selfish players, P1-P2-P3 play the risky action of linking to P4 who then picks the safe action. Players 1, 2, 3 earn 30, and Player 4 earns 24, the lowest in the group. If P4 instead deviates to any risky action, everyone earns 10.

The last mover, P4, could choose to make or break the SPNE outcome: A P4 who has *standard* preferences of maximizing her expected payoff will accept her part in forming the SPNE outcome.[1] But, a P4 who sufficiently dislikes being the lowest paid player would sacrifice her own payments to deviate to a risky action, whereby everyone gets 10. We say such a P4 has *deviant* preferences. Thus, players in P1, P2, and P3 roles experience first-order uncertainty about P4's preferences (standard versus deviant). We use the SSD game to ask our central research question: Are P1's first-order beliefs (what P1 thinks about P4's preferences) and higher-order beliefs (what P1 thinks P2 thinks about P4's preferences, and so on) consistent? Player Pi's choices in SSD reveal her first-order beliefs. The higher the probability Pi assigns to others being a standard type, the higher is her expected payoff from the risky action of linking to others. When this probability is sufficiently large, such that the risky action provides a higher expected payoff than the safe action, we say that Pi is first-order optimistic (O-type); otherwise, we say that Pi is first-order pessimistic (P-type). Thus, when a Pi links to P4, she reveals her first-order optimism.

First-order optimism does not guarantee equilibrium actions from P1 as her higher-order beliefs also influence her choice. If P1 believes that P2 (or P3) is pessimistic about P4, then she foresees P2 (or P3) taking the safe action and thence P4 best-responding by linking to P2 (or P3). Such a P1 expects to get only 10 from linking to P4, irrespective of her first-order optimism/ pessimism. We call P1s with such beliefs as higher-order pessimistic and identify them as the ones who link to P2 (or P3) instead. Finally, P1 plays the equilibrium action of linking to P4 only when P1 is optimistic herself *and* believes that P2 and P3 are also optimistic. Such a P1 reveals herself as both first-order and higher-order optimistic (OO-type). The raw data from the Baseline treatment shows P1 plays the equilibrium action of linking to P4 in only 16% of rounds.[2] Most P1s take the safe action.

Our identification of first and higher-order beliefs relies on four critical properties of the SSD. First, the safe option in SSD reveals first or higher-order pessimism. Second, the game is sequential and the primary uncertainty is about the last moving player. This sequential design breaks the typical circular chain of higher-order reasoning present in simultaneous move games. Third, in SSD, each player moves once as opposed to the canonical experiments on backward induction where all players move multiple times (centipede games by Rosenthal [1981]) or where the same player moves multiple times (chain store paradox games by Selten [1978]). Backward induction in multi-move games requires strong assumptions about how players revise their beliefs

[1] Just like a selfish responder in the Ultimatum game accepts a smaller share of the pie.

[2] A maximum likelihood estimation procedure (described in detail below) finds that only 7% of subjects are OO-type.

about the expected behavior of someone who has previously deviated from the equilibrium path. For example, if j sees i deviate from the equilibrium path what should she assume about i's conformity to backward induction in her future moves? Similar to Dufwenberg and Van Essen [2018], we bypass this controversy by only allowing players to move once. Fourth, the SSD game tree can be easily modified in additional treatments to disentangle first and higher-order beliefs. For example, our 2D treatment modifies the baseline with the following additional rule: if P1 links to P4 then both P2 and P3 are required to also link to P4.[3] This renders inconsequential P1's higher-order beliefs about what P2 or P3 believe about P4. P1's actions now only reflect her first-order beliefs about P4.

We use three diagnostic treatments to separate the belief-related channels from the following two non-belief-related channels that might also explain P1's deviation from the SPNE: P1 might not be using backwards induction, or, P1 might hold non-selfish preferences (they prefer the safe action over free-riding off the safe action of another). The diagnostic treatments establish that the only a small fraction of P1s are deviating for non-belief-related reasons. The inconsistency of first and higher-order beliefs is the primary driver of non-equilibrium play. Our subject-level maximum likelihood estimation suggests that fully 65% of subjects are O-type but not OO-type; first and higher-order beliefs are different for most subjects.

The level-k [Costa-Gomes and Crawford, 2006, Costa-Gomes et al., 2001, Crawford and Iriberri, 2007a,b] and cognitive hierarchy [Camerer et al., 2004] models are the most frequently applied models in which an agent can have divergent first-order and higher-order beliefs (see Crawford et al. [2013] for a detailed survey). These models categorize agents by the degree of rationality that they exhibit: a rational agent best responds, a second-order rational agent best responds to best responses, and so on. The current paper, instead, assumes higher-order rationality (Assumption 1) to identify higher-order beliefs from actions. In addition, we use the diagnostic treatments to identify and restrict our analysis to only those subjects whose behavior is consistent with higher-order rationality.

Beliefs in a game are unobservable, hence, unless inferred through actions, they must be separately elicited. First and higher-order beliefs are difficult to separately elicit without influencing the underlying interaction (Rutström and Wilcox [2009], Gächter and Renner [2010]) and incentives offered by elicitation mechanisms can

[3]In experimental implementation, in order to avoid creating a focal point around linking to P4, the treatment intervention was symmetric with respect to each of P1's risky actions. The simpler description in the main text provides the key intuition of the treatment.

drive false reporting (Danz et al. [2020]). We solve this identification problem by identifying higher-order beliefs directly through actions in a game. Most of the previous experimental literature on beliefs in a game has focused almost exclusively on first-order beliefs: for example, reputation-building under uncertainty about partner's selfish versus reciprocal utility-type has been investigated thoroughly (Andreoni [1993], Cooper et al. [1992]). Notable exceptions are Bosworth [2017] and the literature on guilt aversion (Charness and Dufwenberg [2006], Ellingsen et al. [2010], Khalmetski et al. [2015]), that have also studied the impact of higher-order beliefs in two-player games (what i thinks j thinks i would do).

Uncertainty about other's beliefs, the main focus of this paper, provides a potential explanation for why experimental results often depart from simplistic theoretical predictions. For example, take the influential Baron and Ferejohn [1989] model of multilateral bargaining used by both economists and political scientists. According to the Baron-Ferejohn bargaining procedure, one member of the group is re-picked at random to propose a budget split until a majority agrees with the split. The theory predicts that in a world of *complete information*, the first proposer would enjoy high bargaining power and successfully explains the real-life advantages of leading a bill to the legislative floor.[4] But the experimental literature (for example, Frechette et al. [2003, 2005]) has routinely found under-exploitation of proposal power. We conjecture that this disparity is partly caused by higher-order belief uncertainty in an environment of incomplete information. Both the Baron and Ferejohn [1989] model and real-life bargaining scenarios operate in a world with knowledge and agreement about other's preferences: the former by assumption, and the latter through pre-negotiation dialogue between all participants. Experimental subjects, thrust into a multilateral bargaining experiment, don't enjoy those privileges, and respond by deviating to a fairer and safer allocation of resources, thus failing to take advantage of proposal power. In fact, Agranov and Tergiman [2014] show that once subjects can communicate (cheap-talk), the share of resources extracted by proposers rises to become more aligned with the theory. In contrast, in two-player environments where only first-order and not higher-order uncertainty is relevant, communication

[4]Agranov and Tergiman [2014] mention that in legislative bargaining, the chairman of the appropriations committee, one of the most powerful committees in the Senate, has often been able to steer a disproportionate amount of funds to his district. For example, NYT reported that when Ted Stevens from Alaska held the position, per capita federal spending in Alaska grew by more than 50 percent, by far the highest in the country and almost double the national average.

has the opposite effect: it decreases the proposer/ dictator's share in both two-player bargaining games (Roth [2020]) and dictator games (Andreoni and Rao [2011]).

I. Experimental Design and Hypotheses

The experiment consists of 4 treatments in total: a Baseline treatment and three diagnostic treatments.

I.1. **Baseline treatment.** Every round, subjects are matched in groups of 4. Subjects of a group, P1, P2, P3, P4 are represented by a red node, labelled Red 1 through Red 4.[5] There is also an inert Blue node with no associated player. Figure I.1 displays a screenshot from the experimental interface.

Every round, each player must form a link to another node. Linking to the Blue node is the *safe action* that guarantees 24 points. Linking to a Red node is a *risky action*: If Pi links to Pj ($j \neq i$), then Pi earns 30 points if Pj links to the Blue node and 10 points if Pj links to a Red node. Decisions are made sequentially, with P1 moving first, followed by P2, P3 and P4, and all prior decisions are displayed to subjects when making a decision. In the example screenshot in figure I.1 it is P3's move, and P3 can observe that P1 and P2 each linked to P4. At the end of each round, subjects were shown a screen that summarized the decisions and outcomes made during that round (Figure A.1). Given the payoff structure, playing the safe action of linking to the Blue node may be interpreted as providing a local public good: anyone directly linked to her can enjoy the benefits from her choice.

With standard assumptions and players who maximize expected utility, the SPNE outcome in the Baseline treatment has each of the first 3 players link to P4, and P4 playing the safe action. P1, P2, P3 earn 30, P4 earns 24 and is the lowest paid player. Our first hypothesis documents the SPNE prediction regarding P1 behavior in the Baseline treatment. The SPNE outcome is formally derived in Appendix C.

Hypothesis 1 (SPNE)**.** *In the baseline treatment P1 links to P4.*

Despite the potential of enjoying the highest possible payoff in the SPNE, a player in the role of P1 might deviate from the equilibrium path if one of the following non-standard, behavioral "channels" holds:

I) **Volunteering motive:** P1 prefers to provide the local public good for everyone else by taking the safe action. Such a P1 deviates to taking the safe action.

[5] Player P1 is represented by the node Red 1, and so on. We can refer to the player and the node interchangeably.

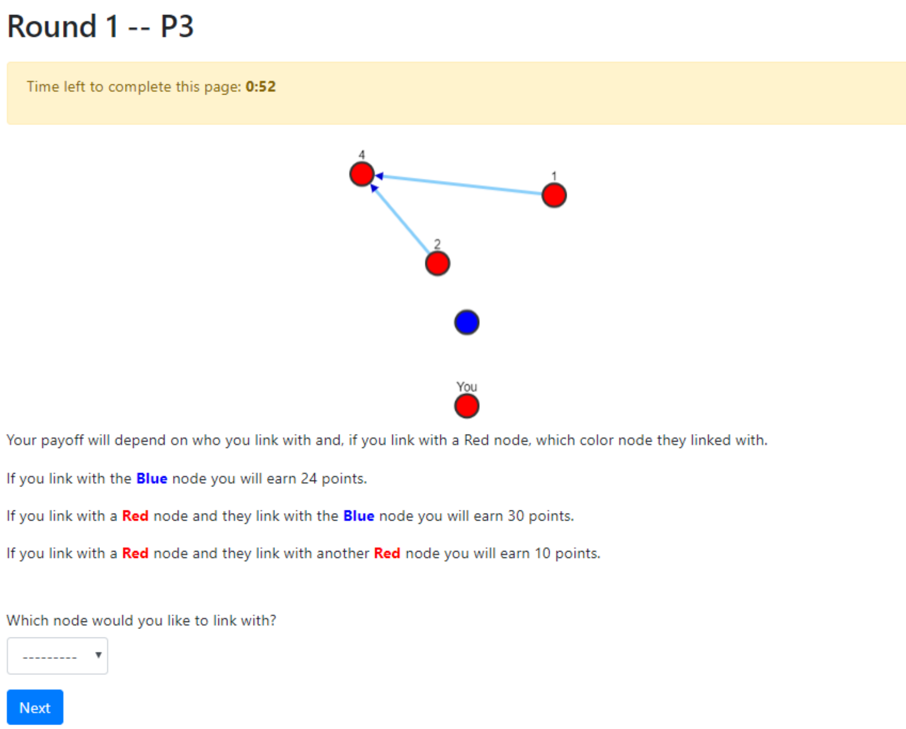


FIGURE I.1. A screenshot from the experimental interface, from the perspective of P3 (Red 3).

II) **Failure of sequential reasoning:** P1 does not reason sequentially about the (SPNE) outcome.[6] Such a P1 deviates to taking the safe action that guarantees a medium payoff.[7]

III) **Higher-order pessimism:** P1 believes that P2 (or P3) are pessimistic, that is, P1 believes that P2 (or P3) believe that P4 would deviate on the equilibrium path, and P2 (or P3) deviates to the safe action in anticipation. Every following player, including P4, would thereafter link to P2 (or P3) and thus linking to P4 only provides a payoff of 10 to P1. Hence, such a P1 prefers linking to P2 (or P3) directly.

IV) **First-order pessimism and higher-order optimism:** P1 believes that P4 would not pick the safe action along the SPNE path. Additionally, she thinks that P2 and P3 are optimistic and would not take the safe action. Thus, such a P1 would deviate to playing the safe action.

[6]She might be boundedly rational or might believe that others are boundedly rational.

[7]Could such a P1 link to P2/ P3 instead? Perhaps, but they must be expecting to get no less than 24 from that action, which also means that they believe that whoever they are linking is to going to take a safe action. We include this case under channel (IV) instead.

A selfish P1 who reasons sequentially under first *and* higher-order optimism would link to P4. For the sake of completeness, let us enumerate this equilibrium-ensuring channel as **(V)**.

Under channels (I)-(IV), P1 would deviate from the SPNE action and the deviating action, when observed in conjunction with diagnostic treatments described below, can reveal the channel at play.

I.2. **Diagnostic treatments:** We modify the Baseline treatment to create the following three diagnostic treatments. Each diagnostic treatment introduces additional rules to the Baseline treatment that mute one or more of the channels (I)-(IV) mentioned above. Our naming convention is to identify each treatment by the number of subjects who make active decisions along the equilibrium path (i.e. the treatment with three Decision makers along the path is referred to as the 3D treatment). The behavioral predictions discussed below are proved formally in Appendix C.

3D treatment: In the 3D treatment, if any three players link to the same player then the linked player is forced to play the safe action. If the linked player has already moved, then the linked player's action is revised to be the safe action. For example, if P1, P2 and P3 all link to P4, then P4 is required to play the safe action. Symmetrically, if P1, P2 and P4 all link to P3, then P3's originally chosen action is revised to be the safe action. The 3D treatment removes any first-order pessimism P1 might have about P4 (channel IV), and any higher-order pessimism that P1 thinks P2/ P3 has about P4 (channel III). Thus, a self-interested P1 who can reason sequentially should not play the safe action.[8]

2D treatment: In the 2D treatment, if P1 plays the safe action then there are no restrictions and the round proceeds as in the Baseline treatment. But, if P1 links to Pj then the other two players, Pk for $k \notin \{1, j\}$, are required to link to Pj. For example, if P1 links to P4, then it becomes a 2-player game between P1 and P4, as both P2 and P3 are automatically also linked to P4. Thus, 2D removes the two middlemen between P1 and P4 on the equilibrium path, and hence removes P1's worry about what P2/ P3 think about P4 on the equilibrium path (channel III). It also removes the difficulty of multiple steps of sequential reasoning (channel II) as

[8]Sequential reasoning suggests that if P1 plays any risky action, say to link to Pj, then the remaining three players will either force Pj to provide the public good or force P1 to provide the public good. Thus, P1 must weakly prefer the outcomes associated with playing a risky action.

only one more player moves after every risky action. Any self interested P1 should take a risky action if and only if she is first-order optimistic.[9]

1D treatment: This treatment combines the interventions in the 3D and 2D treatments. If P1 links to Pj then all others are required to link to Pj, and Pj is required to play the safe action. If P1 plays the safe action, then the 3 remaining players are automatically linked to P1. Thus, P1 alone determines which player receives 24 points, while all remaining players earn 30 points. This treatment removes channels II, III, and IV. Any self interested player would always take a risky action.

Given each diagnostic treatment removes one or multiple mechanisms that causes P1 to choose the safe action, we arrive at the following qualitative hypothesis:

Hypothesis 2 (Behavioral Channels)**.** *P1 takes the safe action in the Baseline treatment more frequently than in any of the 3D, 2D or 1D treatments. 1D has the fewest instances of P1 taking the safe action.*

By comparing how often P1 chooses the safe action in each treatment, we can identify the importance of the individual or pairs of mechanisms. Table 1 describes how P1 would behave across treatments under each possible mechanism I to V. When P1 is indifferent between all risky actions, for example in the 1D treatment, we predict her action as risky. When P1 intends to play a unique risky action, for example under (V) in Baseline, we specify only that specific risky action.

I.3. **Sessions.** Each session contained 12 subjects, randomly and anonymously matched into three groups of 4 subjects each round. Sessions lasted 48 rounds, grouped into 6 blocks of 8 rounds each. Every session contained two out of the four treatments and the two treatments were alternated after every block. Thus, the first, third, fifth blocks ran the first treatment and second, fourth, sixth blocks ran the second treatment. Among two treatments run in the same session the one that generated the smallest game tree was run second. For example, the 1D was always the second treatment, Baseline was always the first treatment, and 3D came before 2D.[10]

We discard the data from the first two blocks (16 rounds) and report only the data from the final 4 blocks to reduce any learning effects that might be caused by subjects

[9]In the 2D treatment, there are three SPNE outcomes for selfish players. In each case, P1 chooses a risky action, say linking to Pj. The best response for Pj is then play the safe action. Similarly, in the 1D treatment there are three equilibrium outcomes where, again, P1 chooses any of the three risky actions.

[10]This ordering was chosen to prevent a focal point in the simpler game "transferring" to the more complex Baseline game.

	Treatments			
Channel	Baseline	3D	2D	1D
I. Volunteering motive	safe	safe	safe	safe
II. Failure of sequential reasoning	safe	safe	risky	risky
III. Higher-order pessimism	P2/ P3	risky	risky	risky
IV. First-order pessimism + higher-order optimism	safe	risky	safe	risky
V. First-order optimism + higher-order optimism	P4	risky	risky	risky

TABLE 1. The mapping between P1's actions and the 5 channels. Linking to Blue is the safe action, linking to P2/ P3/ P4 is the risky action.

seeing a particular treatment first.[11] Subjects were paid for the sum of points earned during one randomly selected block of rounds. Points were converted to Australian Dollars at an exchange rate of $0.15 per point. This implies that payoffs were $4.50, $3.60 and $1.50 per round for the three feasible outcomes. In addition, subjects received a $5 show-up fee and a bonus of up to $3 for comprehension quizzes (up to $1.50 per quiz) that were conducted immediately prior to the start of rounds 1 and 9. Average payments were $40.12 Australian Dollars for sessions that typically lasted between 60 and 90 minutes.[12]

We conducted 12 sessions in total, with a total of 144 subjects. The Baseline, 3D and 2D treatments appeared in 7 sessions each, while the simpler 1D treatment appeared in 3 sessions. We balanced the composition of treatments within sessions, and exploit the overlapping nature of the between-subject portion of the design in our empirical analysis (Table 2).

Our experiments were conducted at the Australian National University (ANU), using student subjects, during Semester 2, 2020 and Semester 1, 2021. Because of both Government and University restrictions on in-person gatherings, we conducted the

[11]The data confirms that there was substantial learning when including all blocks, but that behavior was stable once the first two blocks were dropped.

[12]Sessions that contained the 1D treatment were substantially quicker than others, given that only one subject makes a decision in each round of that treatment. The Baseline treatment was the slowest, as every subject was required to make a decision in every round.

	3D	2D	1D
Baseline	3	3	1
3D		3	1
2D			1

TABLE 2. Number of sessions conducted with each pairwise combination of treatments. For any entry, the row-treatment was run first.

experiments online. In Appendix A we describe the deviations from standard laboratory protocols that were necessary to facilitate the online sessions. Because of the online delivery, the instructions were presented to subjects via an online-lecture style slideshow rather than as traditional text instructions. Pilot sessions found that subjects were more attentive when the slideshow instructions were used. The instruction presentations are available via the appendix.

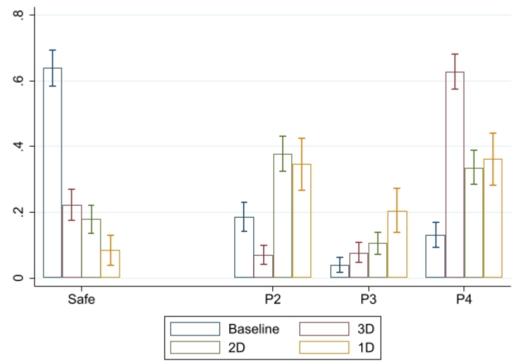
II. RESULTS

Our subjects understood the game and were attentive throughout.[13] We confirm this byincluding results from the comprehension quiz and attention checks in Appendix B. Focusing on our main hypotheses, Hypothesis 1 states that, in the SPNE, P1 in the baseline treatment would link to P4. This hypothesis is rejected: As shown in Figure II.1a, the modal action instead was P1 playing the safe action (taken under channels I, II, or III), followed by P1 linking to P2/ P3 (taken under channel IV) followed by P1 linking to P4 (taken under channel V).

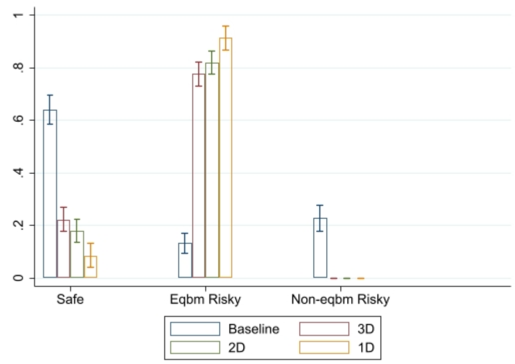
To map actions back into the behavioral channels, we plot the frequency of Safe, Equilibrium risky and Non-equilibrium risky actions in Figure II.1b. The differences are statistically significant. Roughly, the Baseline data implies that there are more P1s with "Higher-order pessimism" (Link to P2/ P3), than with "First and Higher-order optimism" (link to P4).

Hypothesis 2 states that P1 chooses the safe option most often in the Baseline treatment and least often in the 1D treatment. Figures II.1a and II.1b show that P1's behavior shifts away from the safe action in the Baseline treatment towards the risky actions in the diagnostic treatments, and that P1 chooses the safe action

[13]Maintaining the attention of subjects is a particular challenge for online experiments, in comparison to traditional laboratory experiments where external distractions are more easily controlled. We discuss the techniques we used to maintain subject attention online in Appendix A.



(A) P1 action by treatment.



(B) P1 action by treatment (simplified categories).

FIGURE II.1. Behavior across Baseline, 3D, 2D, 1D (Left to Right)

least often in the 1D treatment, in concordance with the hypothesis. The difference in P1 behavior between the Baseline and each of the three diagnostic treatments is statistically significant at the 5% level ($p = 0.002$ for Baseline vs 3D; $p < 0.001$ for Baseline vs 2D; $p = 0.0167$ for Baseline vs 1D), treating each session-treatment pair as on observation.[14] The difference in behavior between the 1D treatment and the other diagnostic treatments is under-powered and not statistically significant ($p = 0.2167$ for 1D vs 3D; $p = 0.5167$ for 1D vs 2D), although this is unsurprising given that we only observe the 1D treatment in three sessions.[15]

The proportion of safe actions across treatments reveals the mechanisms at play. For example, more than 60% of P1s take the safe action in the Baseline treatment. Volunteering, the failure to reason sequentially, and first-order pessimism can together explain this 60%. The relative contribution of each channel to safety-seeking behavior can be disentangled using the behavior of P1 in the diagnostic treatments.

[14]For these p-value calculations we, conservatively, treat each session-treatment pair as on observation. When testing Baseline against F1 or F2, we therefore have 3 sets of paired observations (the three sessions which contained both treatments) and 8 independent observations (four for each treatment). We implement a non-parametric test, without an assumption of equal variances, as outlined in Derrick et al. [2020]. The underlying intuition is that the test statistic is similar to a Mann-Whitney-Wilcoxon statistic, assuming independent samples, that is then adjusted to account for the correlation across the paired observations. For the test of Baseline against F3, we instead implement a Mann-Whitney-Wilcoxon test given that there is only one paired observation.

[15]If we, instead, run a regression analysis that treats the group-round, or the subject, as the level of observation then we do find significant results.

III. Type estimation

We present a very simple model and typology that allow the identification of the behavioral channels that were described in section I. We relegate the technical details to Appendix C. The typology exercise focuses on the behavior of P1.

III.1. **A simple model for typology: Main and auxiliary types.** It is not possible to identify higher-order beliefs for subjects whose play is dictated by volunteering motives (channel I) or failure to reason sequentially (channel II). We therefore classify these two types as auxiliary types: we identify them and explicitly exclude them from the model that identifies higher-order beliefs. P1s whose behavior is dictated by the belief-channels are classified as one of our four main types.

We say that a player Pi *free-rides* on Pj if Pi has linked to Pj and Pj has taken the safe action. We will assume that P1 envisions others having one of two preferences: the standard payoff-maximizing preferences, or deviant preferences under which they would sacrifice payoffs to avoid being free-ridden on when taking the final move.

Definition 1. A subject has Standard (S) preferences if her utility depends only on her own payoff. A subject has Deviant (D) preferences if her preferences resemble S, with one exception: when acting as the final mover she prefers to play a risky action, instead of the safe action, whenever others have linked to her.

Definition 2. A subject is SD-rational if their actions maximize S or D preferences. A subject is **second-order rational** if they are SD-rational and best-respond to the belief that others are SD-rational. A subject is **third-order rational** if they are SD-rational and and best-respond to the belief that others are second-order rational. A subject is **fourth-order rational** if they are SD-rational and and best-respond to the belief that others are third-order rational.

We assume,

Assumption 1 (A1). *Players in P1 role satisfy fourth-order rationality.*

Thus, higher-order rationality implies rationality (with respect to S or D preferences) and the iterated belief that others are also similarly rational. The most natural violation of A1 in our setting would happen for the P1s who have preferences for voluntarily taking the safe action. We screen for such violations of A1 using the 1D treatment. 1D ensures that P1 can free-ride on others at her will.[16] A subject in

[16]Recall that in the 1D treatment that P1 is the only player to move. If P1 chooses the safe action, then all remaining players are required to link to P1. If P1 chooses a risky action, then the player

the role of P1 who yet chooses the safe action violates both S and D preferences: we classify these as the Voluntary or V-type. This is the first auxiliary type for who we cannot identify first and higher-order beliefs from their actions.

P1s who don't reason sequentially as a fourth-order rational player would also violate A1. From the non-Voluntary type P1s, we also screen out such P1s using the 3D treatment. The 3D treatment removes first and higher-order pessimism about a fourth player's action.[17] Irrespective of P1's first or higher-order pessimism, a fourth-order rational P1 would deduce that any risky action by her must result in either her getting 30 or 24, depending on the SPNE played.[18] Such a P1 would weakly prefer playing one of the risky actions. Thus, if a P1 takes the safe action instead, without being a V-type, then she must not satisfy fourth-order rationality. We classify these subjects as non-backwards induction, or NBI type subjects.

A1 explains how our exercise is inherently different from that in Level-k papers, for example Kneeland [2016], that focus on estimating the proportion of subjects satisfying higher-order rationality. We do not pursue identifying different levels of higher-order beliefs. We use a single classification treatment (3D) to screen for higher-order rationality (sequential reasoning), as higher-order rationality is necessary for P1's actions to reveal her higher-order beliefs.

III.2. **SD-rationality and S versus D preferences.** For another simple test of SD-rationality, we can focus on the occasions where a previous mover has already chosen the safe action. In this case, under SD-rationality, a subject should link to the player who chose the safe action and guarantee the maximum feasible payoff of 30 points for herself. There are 1144 observations where we can test for this type of ordinal rationality, and in 1125 of those observations (98.3%) the subject made the SD-rational choice. Of the remaining observations 11 subjects chose the safe action (earning a guaranteed 24 points), 6 subjects linked to a player that had already chosen a risky action (earning 10 points), and 2 subjects linked to a player who had yet to move (generating a possibility of receiving either 10 or 30 points). We interpret this

P1 linked to is required to play the safe action and the remaining two players are required to imitate P1's action.

[17]Recall that the 3D treatment is the Baseline treatment with the following additional rule. If any set of three players all link to the same fourth player, then the fourth player is required to play the safe action. If the fourth player has already chosen a conflicting action, the fourth player's original action is overruled.

[18]This preference will be strict except for the knife-edge case where P1 believes with certainty that the others will always select the equilibrium where P1 receives the lowest payoff.

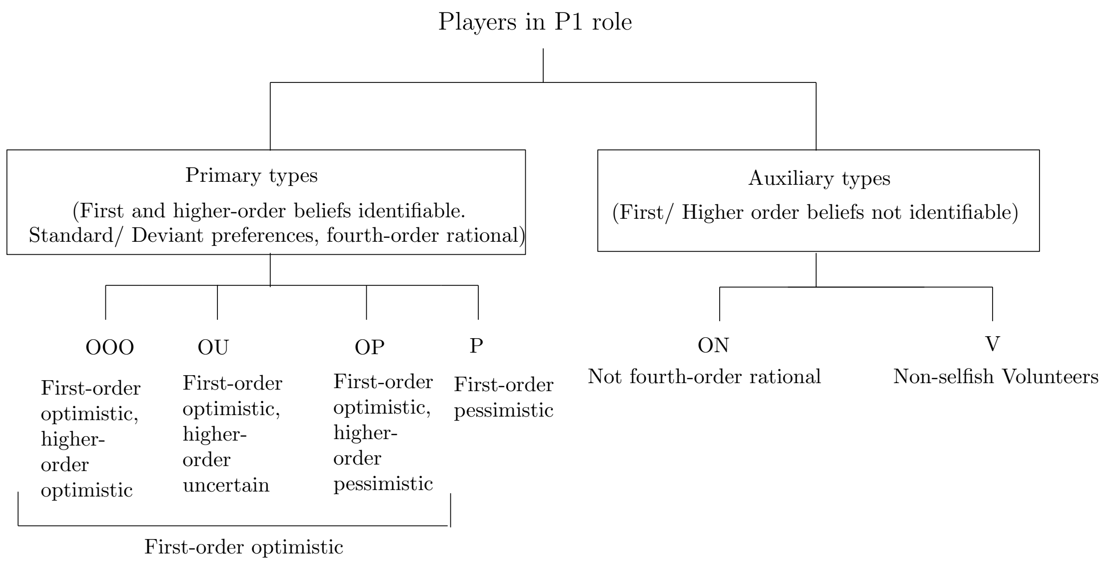


FIGURE III.1. Classification for P1 types

data as justifying our decision to focus our attention on SD-rational types who don't shy away from free-riding.

What proportion of our S-D rational subject pool is consistent with S versus D preferences? We identify this from information sets where S and D types would unambiguously play different actions and find that most of the data is consistent with S preferences. For example, in the 13 rounds which P1, P2 and P3 all link to P4, we observe P4 playing the safe action 100% of the time. Further, in the 2D treatment, P1 links to another player, Pj, 256 times. In each case the decision of Pj is equivalent to the decision facing P4 in the Baseline treatment, and we observe Pj choose the safe action on 248 occasions (96.9% of the time). Given subjects are randomly assigned to player roles, we interpret these numbers as being indicative of almost all subjects having S preferences.

III.3. **Identification strategy for the primary types.** For the subjects who satisfy A1, we can identify the following four primary types (OOO, OU, OP, P) from our treatment variation.

O and P types: We use the 2D treatment to identify P1's (first-order) beliefs about others having either S or D preferences.[19] In the 2D treatment, P1 knows that

[19]Recall that the 2D treatment is the Baseline treatment with the following additional rule. If P1 chooses a risky action, then the two players that P1 did not link to are required to link to the same player that P1 linked to. Thus, if P1 does not play the safe action then the 2D treatment is effectively a two player game between P1 and the player that P1 links to.

if she links to Pj, then the other two players will also automatically link to Pj. Thus P1's higher-order beliefs are irrelevant to P1's decision. P1's decisions depend solely on her first-order optimism/pessimism about Pj having S or D preferences. If P1 links to Pj then P1 reveals herself as the Optimistic type (type O). Instead if P1 plays the safe action, P1 reveals Pessimism and we call this type P.

OP types: Among the first-order optimistic types (identified above), these are the types who are higher order pessimistic. As explained before, if P1 is close to certain that among P2 and P3, the latter is pessimistic about P4, then P1 would link to P3 in Baseline, anticipating that P3 would deviate from the SPNE path to the safe action. Similarly, if P1 is close to certain that P2 is pessimistic about P3 or P4 taking the safe action, then P1 would link to P2 instead in Baseline, anticipating P2's deviation to the safe action.

OOO types: When an optimistic P1 is also close to certain that (i) P2 and P3 are optimistic about P4's play, and (ii) believes that P2 believes that P3 is optimistic about P4's play, then such a P1 would link to P4. We call this the OOO type, as she is optimistic in both first and higher-order beliefs. Thus, the proportion of P1s who link to P4 in the Baseline treatment identifies the OOO types in the population.

OU types: An Optimistic P1 subject who does not assign high enough probabilities to P2 or P3 being either Pessimistic or Optimistic, exhibits second-order Uncertainty. This type of P1 subject doesn't have enough conviction about her higher-order beliefs and is referred to as the OU type. The OU type will play the safe action in the Baseline treatment, preferring the certainty of receiving 24 points over the uncertainty regarding which of P2, P3 or P4 might eventually play the safe action.

In Appendix C we formally establish that the OOO, OP and OU types partition the O type and prove the optimal strategy for each type of P1 in each treatment. The optimal strategies are displayed in Table 3. As can seen in the table, each type has a distinct behavioral profile across the four treatments, allowing identification.

III.4. **Maximum likelihood type estimation.** It is possible to estimate the proportions of types in our sample using the raw data on the distribution of P1 actions across treatments (Figure II.1a). But, this approach would throw away a substantial amount of information. First, because the raw data fails to control for the between-subject aspect of the design, wherein each subject participated in exactly two of the four treatments (see Table 2). Second, because groups are randomly rematched each round we do not observe a balanced panel of subject-level observations in the role of P1, and the raw data might, by chance, over (or under) sample some types of subjects.

		Baseline	3D	2D	1D
Primary types	OOO	link to P4	Risky	Risky	Risky
	OP	link to P2 or P3	Risky	Risky	Risky
	OU	Safe	Risky	Risky	Risky
	PO	Safe	Risky	Safe	Risky
	PP	link to P2 or P3	Risky	Safe	Risky
Auxiliary types	NBI	Safe	Safe	Risky	Risky
	V	Safe	Safe	Safe	Safe

TABLE 3. The set of P1 actions that are feasible for each type of subject, by treatment.

We instead estimate the proportion of types in our data with a mixture model using maximum likelihood estimation, which performs better on both these dimensions.

Our estimation procedure returns estimates of the proportion of each type in our population (π_{OOO},$\pi_{OP}, \pi_{OU}, \pi_P, \pi_{NBI}$ and π_V) and an estimate of the error rate (ϵ). The error rate is defined as the propensity of a type t subject to make a "mistake" and play an action that is inconsistent with type t behavior. For each subject, and each type, we aggregate two values from the raw data: $C_{i,t}$ is the number of observations where subject i, playing as P1, chose an action that was consistent, as outlined in Table 3, with type t; $I_{i,t}$ is the number of observations where subject i, playing as P1, chose an action that was inconsistent with type t. The likelihood of observing $C_{i,t}$ and $I_{i,t}$ for subject i,conditional on the distribution of types, is given by

$$l_i = \sum_{t \in \{OOO,OP,OU,P,NBI,V\}} \pi_t (1-\epsilon)^{C_{i,t}} \epsilon^{I_{i,t}}$$

and the aggregate log likelihood function is then given by

$$LL = \sum_i \log(l_i).$$

The maximum likelihood estimates of the proportions of each type, and the error rate, are presented in Table 4, along with bootstrapped 95% confidence intervals. The confidence intervals are calculated using the bias-corrected and accelerated bootstrap method of Efron [1987], and bootstraps are sampled at the subject level.

	MLE	95% CI
π_{OOO}	0.068	$[0.021, 0.113]$
π_{OP}	0.201	$[0.128, 0.275]$
π_{OU}	0.450	$[0.356, 0.580]$
π_{P}	0.129	$[0.057, 0.189]$
π_{NBI}	0.101	$[0.000, 0.168]$
π_{V}	0.052	$[0.000, 0.124]$
ϵ	0.093	$[0.075, 0.117]$

Table 4. Maximum likelihood estimates and 95% confidence intervals for the proportions of types and the error rate. We report (PO+PP) together under the P type, as the PP was estimated as zero.

The results presented in Table 4 are stark. We estimate that only 15% of subjects fail to satisfy our key identifying assumptions, and are therefore classified as one of two auxiliary types. Strikingly, only 7% of subjects have consistent, optimistic, first-order and higher-order beliefs while almost two-thirds of subjects have optimistic first-order beliefs but do not hold optimistic higher-order beliefs. This is remarkably strong evidence of a systematic inconsistency between first-order and higher-order beliefs.

IV. Discussion

The Baseline treatment of the SSD game is a sequential game with observable actions and a unique SPNE. In the Baseline treatment, we observe only a tiny fraction of groups play the equilibrium strategy. Our experimental design, using three diagnostic treatments, allows us to pinpoint the reason for the failure of equilibrium predictions in this game. At least as importantly, we are able to rule out some common explanations as being substantial drivers of non-equilibrium play.

Our subjects are overwhelmingly rational: over 98% of observations that can be tested for ordinal rationality[20] are in fact rational. Only a small fraction of subjects (the 5% identified as type V subjects) are deviating because of altruistic preferences consistent with warm-glow altruism. Approximately one-tenth of subjects (the 10% identified as type ON subjects) play the SPNE when the game is simplified to a

[20]Link to a previous mover who has already chosen the safe action.

two-player sequential game, but not when faced with the original four-player sequential game, suggesting a cognitive burden associated with longer chains of backwards induction reasoning.

The majority of subjects are deviating from SPNE play because of inconsistent first-order and higher-beliefs. We estimate that, conditional on being able to perform backwards induction and not having altruistic preferences, 85% of our sample believe that P4 will play the safe action with a sufficiently high probability to justify linking to P4 in a 2D treatment.[21] We find evidence that 91% of this subsample or 65% of all subjects, do not hold consistent first-order and higher-order beliefs. Thus, instead of the SPNE action, we observe two modal types of behavior in the Baseline. First, OU types who exhibit uncertainty about the behavior of intermediate movers and therefore play the safe action immediately as P1. Second, OP types who believe that either P2 or P3 are pessimistic (and will therefore play the safe action). The OP types best respond to this belief by linking to either P2 or P3.

Our data shows 97% of decisions by the last mover are selfish rather than deviant. If first-order beliefs were equal to the empirical distribution of actions, then all subjects would be first-order optimistic. We find that a super majority of subjects are indeed optimistic. We observe only 13% of type P subjects. Thus, we find support for the assumption that first-order beliefs are well calibrated to observed behavior, but we do not find support for the assumption that second-order beliefs are consistent with either first-order beliefs or observed behavior.

V. Conclusion

We introduce a novel technique to identify the divergence of first-order and higher-order beliefs in an experimental setting. Our design uses, as a Baseline, a four player sequential social dilemma. We then introduce three variants of the game by pruning the game tree of the Baseline game. The pruning in each variant restricts the beliefs that a player might hold about how others will behave at later nodes in the tree. By comparing behavior across variants, we can make inferences about the structure of beliefs in the Baseline game. Our results show that a common assumption about beliefs, that first-order and higher-order beliefs are consistent, does not hold for a

[21]This figure is the proportion of subjects, conditional on not being V or NBI types, who are OOO, OP or OU types. These subjects can perform backwards induction: they do not play the safe action in the 3D treatment. They also believe that P4 will provide the public good: they do not play the safe action in the 2D treatment.

large proportion of subjects. While we establish this finding in a sequential multi-player game that aids identification, we believe that the conclusion can have material consequences for many games played in the lab or in the field.

Appendix A. Protocol for Online experiment:

Ten minutes prior to the scheduled session start time our subject management software, provided by Sona Systems, distributed a Zoom link to all registered participants. We requested that subjects join the call via a computer, rather than a mobile device. Once subjects joined the Zoom call, we sent each subject a personalized URL via which they could access the experiment. The experiment was programmed using oTree and hosted via oTreeHub, which facilitated this method of delivery [Chen et al., 2016]. Once all subjects had completed the online consent forms, the session began with the experimenter presenting the instructions by sharing his screen and presenting a slideshow to the subjects.[22] Subjects were placed on mute for the duration of the experiment, and Zoom's inbuilt chat function was restricted so that subjects could only message the experimenter (and not each other). Subjects were able to type questions to the experimenter (but could not communicate with each other), who then read the questions out loud to the entire session and provided an answer to the question; the experimenter declined to read or answer the handful of questions that asked for advice about how to best play the game. We tested for subject comprehension with an incentivized quiz, and quiz performance indicated a strong level of understanding of the experimental interface and structure (see Section II for details).

Given that our experimental design required sessions of 12 subjects to advance through the experiment synchronously the possibility of disconnections, or subjects trying to simultaneously complete other tasks, had the potential to slow down, disrupt, or invalidate an entire session of data. We used a four stage mechanism to guarantee the integrity of our data. First, all decision making rounds had a timer. The timer was chosen to be long enough that subjects could make a reasoned decision, but short enough that they were required to maintain attention in order to

[22]We used pilot sessions to fine tune the method of delivery for the instructions. In the initial pilot sessions, we used text instructions that were formatted in a similar fashion to typical instructions used in an in-person experiment. The instructions were then read aloud by the experimenter over the Zoom call, in the standard manner. We found that this method of delivery generated a poor level of attention from our subjects, and thus switched to a more dynamic "online lecture" style format by using a slide deck. After the initial pilot sessions we also sought a change in protocols from the ANU human ethics committee to allow us to request that subjects turn their cameras on during the experiment. While not all subjects were able to comply with the request, the change in protocol did appear to improve engagement with the instructions.

respond before the timer expired.[23] We used a timer length of 60 seconds for the first 5 periods (while subjects were getting used to the interface and game structure), and then shortened the timer to 30 seconds for the remaining rounds. Second, if the timer elapsed before the subject made a decision then the computer would automatically make a decision on behalf of the subject. The computer was programmed to play the Baseline treatment's SPNE strategy in all situations, but subjects were not informed what decision the computer would make (only that the computer would make *a* decision). Third, if a subject timed out in round n then the three opponents of that subject in round $n + 1$were informed of the timeout in the previous round. The instructions clearly explained this to subjects, and emphasized that if they did not receive a message then the three other members of their group were actively participating (and had not been replaced by a computer player). Fourth, we remove from the data set any group-round observations where a member of the group timed out in either the current or previous rounds. Therefore, in every data point we use, every subject in the group made an active decision and knew that every subject in the group had made an active decision in the previous round. The low rate of timeouts observed provides further confidence in the validity of the data set.[24]

Finally, for the first batch of experiments payments were sent electronically to bank accounts that were already held on file by ANU. University processes meant that some payments were delayed, and some subjects were required to update their details with the university prior to payment.[25] In response to these delays, payments for the second batch of sessions were made via PayPal.

Appendix B. Subject comprehension and attention

There were two, incentivized, comprehension quizzes conducted, immediately prior to rounds 1 and 9.[26] For each quiz, subjects began with 10 points and were penalized

[23]We elected not to include a notification alert or sound when it was the subject's turn to make a decision, reasoning that doing so would encourage subjects to pursue other activities in between rounds. By not having an alert, subjects needed to continue to pay close attention to the experimental page to avoid missing their turn.

[24]We observed 144 subjects make 48 decisions each, for a total of 6912 decisions. Of those 6912 decisions we observed only 119 timeouts (i.e. less than 2% of decisions timed out). For 77 of the 118 timeouts that were not in the final round of a session, the subject did not timeout again in the following round.

[25]Four subjects did not respond to emails requesting that they update their details, and payments were currently unable to be processed for those four subjects.

[26]Recall that each session contained two treatments. Round 9 was the first round in which the subjects played the second treatment.

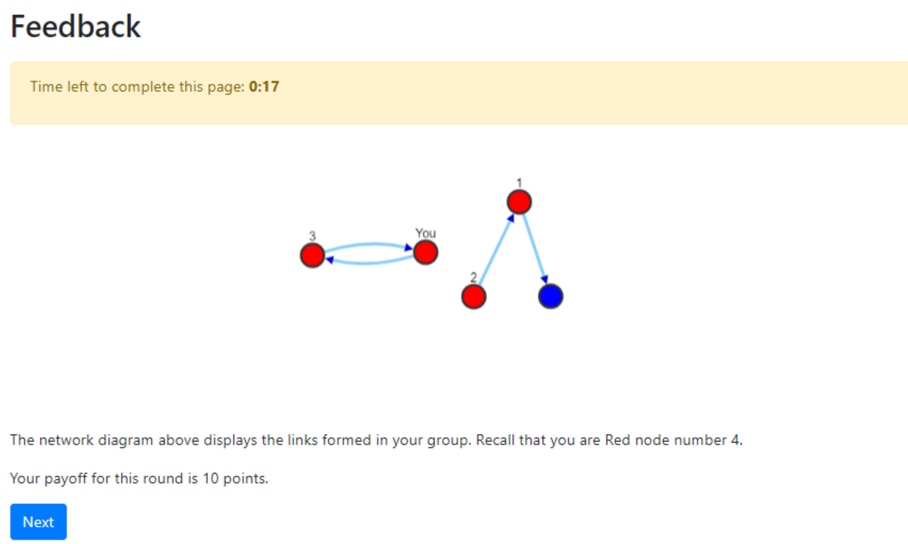
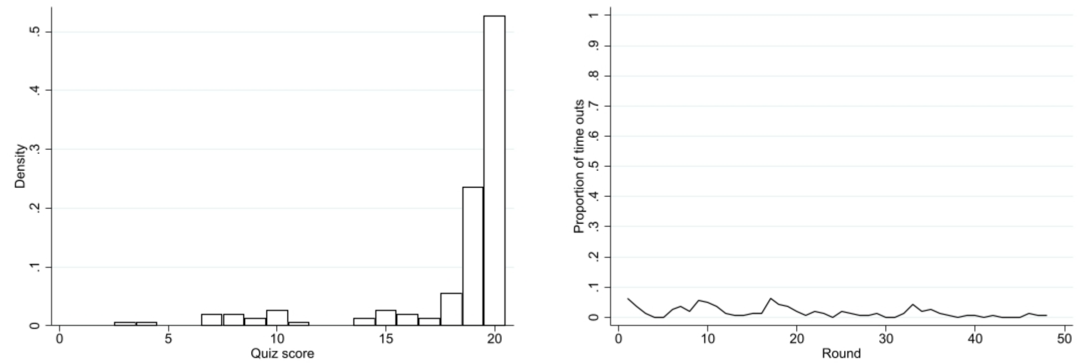


FIGURE A.1. A screenshot of the feedback screen, from the perspective of P4 (Red 4).

one point per mistake made.[27] Subjects had to answer all questions correctly before the experiment continues.[28] The questions asked about the payoffs to players from an already formed network, payoffs to a player who made a hypothetical move in a partially formed network, and whether given networks were feasible or infeasible networks. Overall, subjects displayed an excellent understanding with three-quarters of subjects making one or fewer mistakes across the two quizzes. Figure B.1a shows the histogram of aggregate performance across the two quizzes.

We also track the attention subjects paid to the experiment by examining the proportion of time outs. The timer was 60 seconds for the first 5 rounds, reduced to 30 seconds for the remaining rounds – short enough that an inattentive subject (e.g. checking emails on another web page) was unlikely to respond in time, but long enough that an attentive subject would comfortably make a decision within the allotted time.[29] The proportion of time outs by round is displayed in figure B.1b. Overall, the proportion of time outs is low. There are small spikes in rounds 1, 9, 17

[27]The score for each quiz had a floor of 0 points, so that subjects could not lose money during the quizzes. The points were valued at the same rate as points earned playing the game: $0.15 per point.
[28]If a subject made a mistake, they were prompted to try again until they found the correct answer.
[29]Some timeouts appear to be caused by genuine technical difficulties. For example, one subject dropped out for several rounds while searching for a power point at which to recharge his laptop (which had run out of batteries).



(A) Histogram of quiz-performance (out of 20) across the two comprehension quizzes.

(B) Proportion of subject time outs per round.

FIGURE B.1. Subject comprehension and time-outs.

and 33. Rounds 1 and 9 immediately followed the quizzes, and some subjects might have navigated away from the experiment page while waiting for others to complete the quiz. There were also short breaks between each block of 8 rounds (usually less than 30 seconds) which coincides with the increased time outs in rounds 17 and 33. In the subsequent analysis we remove from the data all groups in which any subject timed out, and all groups in which any subject timed out in the previous round.[30]

APPENDIX C. THEORETICAL RESULTS

C.1. **Uncertainty and related typology.** This section constructs the typology that underlies the intuition presented in the main text. The analysis here assumes Assumption 1. That is, P1 is fourth-order rational.

Let $\Delta(Y)$ denote the space of probabilities on a given set Y, and interpret these probabilities as beliefs. We write Δ_i to specify the belief held by player Pi. Given the sequential structure of the game, P4's belief is always degenerate. Stepping backwards, P3's only payoff-relevant uncertainty is about P4. Thus, P3's belief belongs

[30]As discussed in detail in Appendix A, if a subject times out then the computer automatically makes a decision on the subject's behalf (and continues to do so for future rounds until the subject returns to the experiment). Subjects are informed if any members of their 4-person matching group failed to make a decision in the previous round. Thus, if a subject times out in round t then the behavior of the subject's opponents may be affected in rounds t and $t + 1$.

to the space

$$\underbrace{\Delta_3\left(\underbrace{t_4}_{\text{P4's type}}\right)}_{\text{P3's belief}} \tag{C.1}$$

Similarly, P2 will hold beliefs about P3's first-order belief about P4, and P4 's type. Thus, P2's belief belongs to the space

$$\Delta_2\left(\underbrace{\Delta_3\quad\underbrace{(t_4)}_{\text{P4's type}}}_{\text{P3's belief}} \times \underbrace{t_4}_{\text{P4's type}}\right) \tag{C.2}$$

By defining the beliefs over the product space, we allow P2 to express correlation between different components. Finally, P1's belief belongs to the space

$$\Delta_1\left(\underbrace{\Delta_2\left(\underbrace{\Delta_3\quad\underbrace{(t_4)}_{\text{P4's type}}}_{\text{P3's belief}} \times \underbrace{t_4}_{\text{P4's type}}\right)}_{\text{P2's belief}} \times \underbrace{\Delta_3\quad\underbrace{(t_4)}_{\text{P4's type}}}_{\text{P3's belief}} \times \underbrace{t_4}_{\text{P4's type}}\right) \tag{C.3}$$

The first term in (C.3) is used to define P1's belief about P2's belief, Δ_2, the second term is P1's belief about P3's beliefs (Δ_3) and the final term is P1's belief about P4's type (t_4). Abusing notation, we will use Δ_i^p to denote the probability i's belief (an element of Δ_i) assigns to some event in its domain.

We also impose the tie-breaking assumption that whenever a player is indifferent between two actions, they play the action that is consistent with the SPNE whenever possible and play risky actions over the safe action. While this assumption is not necessary for the analysis, it does simplify the proofs presented below.

Lemma 3. *If Pj plays the safe action then, for all $k > j$, any Pk who is SD-rational links to Pj.*

Proof. If $j = 4$ then the statement is trivially true. We assume $j < 4$. Whenever P4 has the move, and at least one previous mover has played the safe action, P4 links to one of the players who played the safe action. Therefore, when P3 has the move, and at least one previous mover has played the safe action, P3 reasons that linking to P4 will earn 10. Given this, P3 links to one of the players who played the safe action. When P2 has the move, and P1 has played the safe action, P2 reasons that linking to P3 or P4 will earn 10. Given this, P2 links to P1. □

Proposition 4. *(i) P1 is certain that if she or P2 have played the safe action, then P3 links to a player who has played the safe action.*

(ii) P1 is certain that if she or P2 have played risky actions, then there exists a $\bar{q}$ *such that P3 links to P4 if and only if P3's belief satisfies* $\Delta_3^p(\underbrace{S}_{t_4=S}) \geq \bar{q}$*, and P3 plays the safe action if and only if* $\Delta_3^p(\underbrace{S}_{t_4=S}) < \bar{q}$*.*

Proof. (i) Follows directly from Lemma 3.

(ii) Suppose that P1 and P2 have both played a risky action. If P3 chooses the safe action, she earns 24 with certainty. If P3 links to P1 or P2, she earns 10 with certainty. If P3 links to P4, her earnings depend on P4's action. If P4 plays the safe action, which is P4's best response if P4 is of type S, then P3 earns 30. If P4 plays any risky action, which is P4's best response if P4 is of type D, then P3 earns 10.

Therefore, P3 is indifferent between linking to P4 and playing the risky action when $u(30)\Delta_3^p(S) + u(10)(1 - \Delta_3^p(S)) = u(24)$. Set $\bar{q} = \frac{u(24)-u(10)}{u(30)-u(10)}$, and note that monotonicity of the utility function implies that $0 \leq \bar{q} \leq 1$.

If $\Delta_3^p(S) > \bar{q}$ then linking to P4 is the unique best response, and if $\Delta_3^p(S) < \bar{q}$ then the safe action is the unique best response. Further, given the tie breaking assumption, P3 will also link to P4 when $\Delta_3^p(S) = \bar{q}$. Conversely, if P4 is a best response, then $\Delta_3^p(S) \geq \bar{q}$ and if the safe action is best response then $\Delta_3^p(S) < \bar{q}$. □

Henceforth, we denote the set P3's beliefs that satisfy $\Delta_3^p(\underbrace{S}_{t_4=S}) \geq \bar{q}$ by $\Delta_3^* \subseteq \Delta_3$. When P3's beliefs satisfy this condition we say that P3 is Optimistic about the behavior of P4, and otherwise say that P3 is Pessimistic.

Proposition 5. *(i) P1 is certain that after she has played the safe action, P2 links to P1.*

(ii) P1 is certain that after she has played a risky action, P2 links to P4 if and only if P2's beliefs lie in Δ_2^**,*

$$\Delta_2^* = \{\Delta_2^p : \Delta_2^p(\underbrace{\Delta_3^*, t_4 = S}_{\text{P3 is optimistic, P4 is S}}) \geq \bar{q}\} \tag{C.4}$$

(iii) P1 is certain that after she has played a risky action, P2 links to P3 if and only if P2's beliefs lie in Δ_2^{**}*, where*

(C.5)
$$\Delta_2^{**} = \{\Delta_2^p : \Delta_2^p(\underbrace{(\Delta_3^*)^C}_{\text{P3 is pessimistic}}, t_4) \geq \bar{q}\}$$

(iv) P1 is certain that after she has played a risky action, P2 plays the safe action if and only if P2's beliefs lie in $(\Delta_2^* \cup \Delta_2^{**})^C$*.*

Proof. (i) Follows directly from Lemma 3.

(ii) In the event $(\Delta_3^*, t_4 = S)$, P3 will link to P4 (Proposition 4) and P4 will play the safe action. If $\Delta_2^p(\Delta_3^*, t_3, t_4 = S) > \bar{q}$ then linking to P4 is the unique best response. Further, given the tie breaking assumption, P3 will also link to P4 when $\Delta_2^p(\Delta_3^*, t_4 = S) = \bar{q}$. Conversely, if P4 is a best response for the S type, then $\Delta_2^p(\Delta_3^*, t_4 = S) \geq \bar{q}$.

(iii) In the event $((\Delta_3^*)^C, t_4)$, P3 will play the safe action. The remainder of the proof follows the argument in part (ii).

(iv) If $\Delta_2^p(\Delta_3^*, t_4 = S) < \bar{q}$ and $\Delta_2^p((\Delta_3^*)^C, t_4) < \bar{q}$ then playing the safe action is the best response – the expected payoff from linking to either P3 or P4 is less than $u(24)$. Conversely, if playing the safe action is the best response then $\Delta_2^p(\Delta_3^*, t_4 = S) < \bar{q}$ and $\Delta_2^p((\Delta_3^*)^C, t_4) < \bar{q}$. □

Notice that whenever P2 has beliefs that satisfy Δ_2^* then P2 is both Optimistic herself and Optimistic about P3's optimism.[31] Thus, if P2 has beliefs that satisfy Δ_2^* we say that P2 is Optimistic-Optimistic, or first and second order optimistic (given that P2 is optimistic that P4 will play the safe action and that P3 will link to P4).

Proposition 6. *(i) P1 links to P4 if and only if P1's beliefs satisfy*

(C.6)
$$\Delta_1^p\left(\Delta_2^* \cup \Delta_2^{**}, \Delta_3^*, t_4 = S\right) \geq \bar{q}$$

(ii) P1 links to P2 if and only if P1's beliefs satisfy

(C.7)
$$\Delta_1^p\left(\underbrace{(\Delta_2^* \cup \Delta_2^{**})^C}_{\text{P2 is pessimistic, hence plays safe}}, \Delta_3, t_4\right) \geq \bar{q}$$

(iii) P1 links to P3 if and only if P1's beliefs satisfy

[31] Analogous to the case for P3, we can state that P2 is Optimistic if $\Delta_2^p(\Delta_3, t_4 = S) \geq \bar{q}$. That beliefs in Δ_2^* are Optimistic is immediate.

$$\Delta_1^p\left(\Delta_2^* \cup \Delta_2^{**}, (\Delta_3^*)^C, t_4\right) \geq \bar{q} \tag{C.8}$$

(iv) P1 plays the safe action if and only if P1's beliefs do not satisfy the conditions in (i), (ii) or (iii).

Proof. (i) If P1 links to P4 then, in the event $(\Delta_2^* \cup \Delta_2^{**}, \Delta_3^*, t_4 = S)$, P2 and P3 will choose a risky action and P4 will play the safe action.The remainder of the proof follows the argument in Proposition 4 in part (ii).

(ii) If P1 links to P2 then then in the event $\left((\Delta_2^* \cup \Delta_2^{**})^C, \Delta_3, t_3, t_4\right)$ P2 will play the safe action. This is the only case where P2 plays the safe action. The remainder of the proof follows the argument in Proposition 4 in part (ii).

(iii) If P1 links to P2 then in the event $\left(\Delta_2^* \cup \Delta_2^{**}, (\Delta_3^*)^C, t_4\right)$ P2 will play a risky action (Proposition 5), and P3 will play the safe action (Proposition 4). The remainder of the proof follows the argument in Proposition 4 in part (ii).

(iv) The result follows immediately from the results in parts (i), (ii) and (iii). □

When P1's beliefs satisfy equation C.6 we say that P1 is first-, second- and third-order optimistic (OOO): P1 is optimistic that P4 is type S, believes that P3 is optimistic about P4, and believes that P2 is optimistic about P4 and believes that P3 is optimistic about P4. When P1's beliefs satisfy equation C.7 or equation C.8 we say that P1 is higher-order pessimistic: P1 believes that either P3 is pessimistic about P4, or that P2 is pessimistic about P4. Finally, when P1's beliefs satisfy none of equations C.6, C.7 or C.8 we say that P1 is higher-order uncertain: P1 does not have strong enough beliefs to justify playing any risky action.

The following propositions outline the behavior of P1 in the diagnostic treatments. We continue to assume A1 and A2.

Proposition 7. *In the 1D treatment, P1 will always play a risky action.*

Proof. Immediate. Playing the safe action pays 24, and playing any risky action pays 30. □

In the 2D treatment, P1's belief can be summarized by the value of $\Delta_1^p(S)$: the probability that P1 believes that the final mover in the game is of type S. P1 is Optimistic if $\Delta_1^p(S) \geq \bar{q}$ and Pessimistic otherwise.

Proposition 8. *In the 2D treatment, P1 will play a risky action if and only if* $\Delta_1^p(S) \geq \bar{q}$.

Proof. Follows the proof of part (ii) of Proposition 4. □

Finally, in the 3D treatment, the rules of the game force all players to behave as if they are type S. Therefore, applying the same logic as in Proposition 9, the behavior of P1, given A1, is identical to the SPNE. Given the linking rules in the 3D treatment are bespoke, we outline the proof of the SPNE in detail in the following proposition.

Proposition 9. *In the Baseline treatment with S players, the unique Subgame Perfect Nash Equilibrium outcome is for P1 to link to P4, P2 to link to P4, P3 to link to P4, and P4 to take the safe action.*

Proof. We proceed via backwards induction, beginning with P4. At any decision node where no previous mover has played the safe action, P4's best response is to play the safe action. If at least one other player has played the safe action, then P4's best response is to link to a player who has played the safe action.

Next, consider P3. If either P1 or P2 has played the safe action, then P3's best response is to link to a player who played the safe action. In this case P3 earns 30 points, rather than 24 if P3 plays the safe action or 10 if P3 links to P4 (as P4 will link to either P1 or P2). If neither P1 nor P2 has played the safe action, then P3's best response is to link to P4, who will then play the safe action, which pays 30 points for P3.

Next, consider P2. If P1 has played the safe action, then P2's best response is to link to P1. If P1 did not play the safe action, then P2's best response is to link to P4. Doing so earns P2 30 points, given the P3 will link to P4 and P4 will play the safe action, compared to either 24 or 10 for any other action.

Finally, consider P1. Given the behavior of P2, P3 and P4 outlined above, P1 faces the following choices. If P1 plays the safe action, then P1 will earn 24 points. If P1 links to P2 or P3, then P1 will earn 10 points (as both P2 and P3 will link to P4). If P1 links to P4 then P1 will earn 30 points. □

References

Marina Agranov and Chloe Tergiman. Communication in multilateral bargaining. *Journal of Public Economics*, 118:75–85, 2014.

James Andreoni. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *Economic Journal*, 103(418):570–585, 1993.

James Andreoni and Justin M Rao. The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of public economics*, 95(7-8):513–520, 2011.

David P Baron and John A Ferejohn. Bargaining in legislatures. *American political science review*, 83(4):1181–1206, 1989.

Steven J. Bosworth. The importance of higher-order beliefs to successful coordination. *Experimental Economics*, 20:237–258, 2017.

Colin F. Camerer, Tech-Hua Ho, and Juin-Kuan Chong. A cognitive heirarchy model of games. *Quarterly Journal of Economics*, 119(3):861–898, 2004.

Gary Charness and Martin Dufwenberg. Promises and partnership. *Econometrica*, 74(6):1579–1601, 2006.

Daniel L. Chen, Martin Schonger, and Chris Wickens. otree – an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016.

R. Cooper, D.W. DeJong, and T.W. Ross. Cooperation without reputation: Experimental evidence from prisoner's dilemma games. Papers 36, Boston University - Industry Studies Programme, 1992. URL http://ideas.repec.org/p/fth/bostin/36.html.

Miguel Costa-Gomes and Vincent P. Crawford. Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96(5):1737–1768, 2006.

Miguel Costa-Gomes, Vincent P. Crawford, and Bruno Broseta. Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 69(5):1193–1235, 2001.

Vincent P. Crawford and Nagore Iriberri. Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica*, 75(6):1721–1770, 2007a.

Vincent P. Crawford and Nagore Iriberri. Fatal attraction: Salience, naivete, and sophistication in experimental "hide-and-seek" games. *American Economic Review*, 97(5):1731–1750, 2007b.

Vincent P. Crawford, Miguel A. Costa-Gomes, and Nagore Iriberri. Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51(1):5–62, 2013.

David Danz, Lise Vesterlund, and Alistair J Wilson. Belief elicitation: Limiting truth telling with information on incentives. Technical report, National Bureau of

Economic Research, 2020.

Ben Derrick, Paul White, and Deirdre Toher. Parametric and non-parametric tests for the comparison of two samples which both include paired and unpaired observations. *Journal of Modern Applied Statistical Methods*, 18(1), 2020.

Martin Dufwenberg and Matt Van Essen. King of the hill: Giving backward induction its best shot. *Games and Economic Behavior*, 112:125–138, 2018.

Bradley Efron. Better bootstrap confidence intervals. *Journal of the American Statistical Association*, pages 171–185, 1987.

Tore Ellingsen, Magnus Johannesson, Sigve Tjøtta, and Gaute Torsvik. Testing guilt aversion. *Games and Economic Behavior*, 68(1):95–107, 2010.

Guillaume Frechette, John H Kagel, and Massimo Morelli. Nominal bargaining power, selection protocol, and discounting in legislative bargaining. *Journal of Public Economics*, 89(8):1497–1517, 2005.

Guillaume R Frechette, John H Kagel, and Steven F Lehrer. Bargaining in legislatures: An experimental investigation of open versus closed amendment rules. *American Political science review*, 97(2):221–232, 2003.

Simon Gächter and Elke Renner. The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, 13(3):364–377, 2010.

Kiryl Khalmetski, Axel Ockenfels, and Peter Werner. Surprising gifts: Theory and laboratory evidence. *Journal of Economic Theory*, 159:163–208, 2015.

Terri Kneeland. Coordination under limited depth of reasoning. *Games and Economic Behavior*, 96:49–64, 2016.

Robert W Rosenthal. Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic theory*, 25(1):92–100, 1981.

Alvin E Roth. 4. bargaining experiments. In *The handbook of experimental economics*, pages 253–348. Princeton University Press, 2020.

E Elisabet Rutström and Nathaniel T Wilcox. Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior*, 67(2):616–632, 2009.

Reinhard Selten. The chain store paradox. *Theory and decision*, 9(2):127–159, 1978.