Journal of Economic Psychology 108 (2025) 102806

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Heuristic centred-belief players☆

Irenaeus Wolff

Thurgau Institute of Economics (TWI)/University of Konstanz, Hafenstrasse 6, 8280 Kreuzlingen, Switzerland

ARTICLE INFO

JEL classification:
C72
C92
D83

Keywords:
Nash-equilibrium
Quantal-response equilibrium
Level-*k*
Cognitive-hierarchy
Salience theory
Noisy introspection
Central-tendency bias

ABSTRACT

Strategic behaviour often diverges from Nash-equilibrium, in particular in inexperienced play. Studying a class of games in which participants choose their payment and receive it as long as their opponent chooses a different amount, I show that none of the popular models of behavioural game theory predicts the predominant aggregate choice pattern consistently. And yet, noisy introspection (Goeree and Holt, 2004) readily accounts for about half of the individual observations. The reason for the apparent paradox and the mis-match of the aggregate data and the models is a disregarded behavioural type that makes up about 25% of the population. These 25% display a specific form of central-tendency bias, holding beliefs that peak in the centre of the option set and that are roughly symmetric. In addition, the players show a more heuristic process translating their belief into actions, as their choices cannot be explained readily by quantal responding. The behavioural pattern of a 'centred belief' in connection with boundedly-rational decision-making is present also in another prominent game from the literature on behavioural game theory, the 11–20 game. Finally, I show that classifying players as 'heuristic centred-belief types' by one game's beliefs has predictive power for behaviour in the other game.

1. Introduction

This paper experimentally studies a simple game I call the name-your-prize game. The game is a simple discoordination game, so that players only obtain a positive payoff if they choose a different option than their opponent:

> *Please choose one of the following payoffs, [27, 30, or 33]. You receive the payoff you choose, but only if the other player chooses a different payoff than you. Otherwise, you receive nothing.*

The paper proceeds in four steps. First, I observe the aggregate data pattern and establish the empirical pattern's level of robustness by varying features of the game and the experiment in 18 treatments. Second, I contrast the empirical response distribution with popular behavioural-game-theory models and observe that none of the models readily predicts aggregate behaviour. Third, I explore potential reasons for the unexpected pattern by eliciting players' strategies and beliefs. The belief analysis shows that while noisy introspection (Goeree & Holt, 2004) does account for the individual behaviour of slightly more than 50% of the population, the aggregate pattern is driven mainly by a new behavioural type I call *heuristic centred-belief player*. And fourth, I find that this type

☆ I am grateful to all the many people who helped me write this paper, in particular Ankush Asri, David Cooper, Miguel Costa-Gomes, Fabian Dvorak, Urs Fischbacher, Moath Hussien, Andrea Isoni, Max Lobeck, Ariel Rubinstein, Julia Werner, the research group at the Thurgau Institute of Economics, participants of the microeconomics seminar at the University of Konstanz, research seminars at Otto-von-Guericke University Magdeburg and Bremen University, as well as participants of FUR 2022 (Ghent), and the 2022 UEA Annual Workshop on Behavioural Game Theory. Special thanks go to Ariel Rubinstein for implementing the "Rob-Pool" treatments on his website http://gametheory.tau.ac.il/. All data I used is available under http://x-econ.org/xecon/#!Detail/10.23663/x2698, together with .r-files that allow to re-produce the figures and tables of the paper (with the exception of Table 1, which has to be combined from two different files provided in the data archive).

E-mail address: wolff@twi-kreuzlingen.ch.

https://doi.org/10.1016/j.joep.2025.102806
Received 28 September 2023; Received in revised form 24 February 2025; Accepted 25 February 2025
Available online 7 March 2025

exists in other (existing) data sets from the literature, too. Moreover, I show that its identification in one game allows for certain predictions for another game.

Altogether, the paper suggests three things: first, to understand behaviour even in simple games, we need to look at a mix of types, as no single model will explain the data by itself. This does not always have to be a mixture of a behavioural-economics model and a heuristic as in this study (in an updating task, Achtziger & Alós-Ferrer, 2014, have shown that a mixture of rational Bayesian updaters and reinforcement learners explain their data best; in other contexts, researchers might have to consider yet other types of mixes: for example, think of trust games and inequality-aversion and guilt-aversion or any other non-nested models of social preferences). Second, the paper supports noisy introspection for explaining part of the data. Third, for games in which the players' available alternatives have a natural order, the paper identifies *heuristic centred-belief players* as an additional type to watch out for: *Heuristic centred-belief players* can add distinctive 'features' to the aggregate data pattern (the name-your-prize game and the 11–20 game described below are prime examples). A full characterization of the types' heuristic remains a topic for further research. I next provide more context on each of the four steps of the paper.

Step one: The name-your-prize game. The baseline game has a unique symmetric Nash equilibrium under risk-neutrality, in which both players choose 27, 30, and 33 with probabilities of 26%, 34%, and 40%, respectively (risk-aversion would mainly push the equilibrium strategy towards uniformity). The frequencies I observe when running the game with 168 participants are 31%, 21%, and 48%. Thus, the second-highest action is chosen with a much lower frequency than under the Nash-equilibrium, resulting in a U-shaped rather than a monotonic aggregate data pattern.

To test the robustness of the finding, I run a substantial number of additional treatments: first, I change the prizes to have an uneven number in the middle (rather than "30") and to explore larger differences. Then, I vary whether all players are paid and whether the game is only an additional part of another experiment or the first (and in some treatments: only) part. In yet other treatments, I vary the number of alternatives and the participant pool. And finally, I add two versions in which participants simultaneously play a whole population of other participants.

In all 18 treatments, the second-highest prize is chosen less often than in the respective symmetric Nash equilibrium. Moreover, in 13 of the 18 experiments, the modal choice is the highest prize, but the second-highest prize is chosen less often than the third-highest prize. In a 14th case (a variant with 5 options), it is the third-highest prize that is chosen less often than the fourth-highest prize. The remaining 4 treatments show monotonic patterns.

Step two: Popular models of behavioural game theory. Deviations from Nash-equilibrium should come as no surprise in a one-shot environment. In case of such a deviation, however, we should expect the behaviour to be captured by at least one of the popular alternative models suggested by behavioural economics. And yet, note that the standard quantal-response-equilibrium, cognitive-hierarchy, level-k, salience-theory, and impulse-balance-equilibrium models also predict monotonic patterns (or, in some cases and under certain parameters, even a hump-shaped pattern). Team reasoning does not lead to a different prediction than the symmetric Nash equilibrium, either, because there is no way of distinguishing the players, and so, players cannot coordinate their choices. Noisy introspection (Goeree & Holt, 2004) is the only model to fit the predominant aggregate data pattern in a minority of the variations displaying the pattern (5 out of 14; the baseline game being one of them).

Step three: Identifying the new heuristic. I run a treatment with an incentivized elicitation of participants' reasoning and a belief-elicitation treatment. In the end, a cluster analysis of the beliefs data provides the key information to understand the game data. Roughly half of the participants can be classified as conforming to noisy introspection, the other half cannot. Among this second half, there is a robustly-classified cluster of about 25% of the whole population who exhibit a specific form of the central-tendency bias (see, for example, Crosetto et al., 2020, and references cited therein). These 'heuristic centred-belief players' show roughly symmetric (though somewhat left-skewed) beliefs that put the highest probability mass on the opponent's 'central' option(s). They choose "30" particularly rarely, and excluding them from the data would result in a monotonic aggregate data pattern.

Step four: Heuristic centred-belief players and the 11–20 game. Having identified the heuristic centred-belief type, I set out to explore whether the type was particular to the game(s) I had been studying. Thus, I looked for a data set of a game that is (i) prominent in the experimental literature on strategic behaviour, and (ii) would include elicitation of probabilistic beliefs, so that the data set would allow the identification of heuristic centred-belief players. The data set on the 11–20 game provided by Goeree et al. (2018) fulfilled both conditions. It includes an elicitation of probabilistic beliefs, and the game is subject to an extended debate about whether the initially-proposed level-k explanation for the data is a sufficiently good approximation (see, for example, Alós-Ferrer & Buckenmaier, 2021, and references cited therein).

In the 11–20 game, players obtain the amount they request between 11 and 20 currency units, plus a bonus of 20 units if they request exactly 1 unit less than their opponent. There are centred-belief players in the standard treatment in which the order of options is 11–12–...–19–20, but not in the treatments with perturbed orders (19–18–17–...—12–11–20 and 14–13–12–11–19–18–17–16–15–20). As in my discoordination games, the centred-belief players differ from other players in both their (roughly symmetric and hump-shaped) beliefs and their choices also in the 11–20 game. In this game, centred-belief players account for 50%–67% of all choices below 17, even though they are only 8%–17% of the population.

In a final treatment variation, I test whether heuristic centred-belief play is a strategy that people sometimes adopt when playing games like the ones studied here, or whether heuristic centred-belief players are a type of person who routinely uses such a strategy across situations. In particular, I let participants play a name-your-prize game, ask for their beliefs in that game, and then let them play the 11–20 game. I run the same cluster analysis as before on the name-your-prize beliefs, categorizing participants into centred-belief players and other players, and then look at the two groups' behaviour in the 11–20 game.

The centred-belief players again are much more likely to choose an action below 17 than all other player types, in this case, by a factor of three. Hence, I again find a fraction of roughly 25% heuristic centred-belief players (who produce a non-monotonic pattern, choosing the low, intermediate, and high payoffs with relative frequencies of 28%, 12%, and 60%, respectively). More importantly, however, heuristic centred-belief play seems to be a strategy that a certain part of the population applies *across* games.

2. Related literature

This paper contributes to three strands of literature. First, there is a large and growing literature on strategic reasoning in one-shot games (for an introduction, *cf.*, Crawford et al., 2013). In most of this literature, there is a particular model of behavioural game theory that explains behaviour best, but the best-performing model differs between papers (and sometimes, even within papers, as in Bardsley et al., 2010). In this sense, my paper falls into the same category as, for example, Goeree et al. (2018), showing that noisy introspection accounts for the behaviour of a substantial part (one-half) of the population particularly well.

However, my paper also shows that participants' aggregate response pattern (and a non-negligible sub-population's individual response patterns) cannot be explained readily by any of the popular models of behavioural game theory. Instead, behaviour is governed by a more heuristic decision-making procedure. In this sense, the paper is related to the huge literature following the works of Simon (1947) and Tversky and Kahneman (1974), showing that in certain situations, simple decision rules explain behaviour best. Recent examples that also have a strong coordination aspect are Greiner (2023) and De Kwaadsteniet et al. (2023). Greiner (2023) shows that people's bargaining behaviour cannot be explained by beliefs consistent with their strategies. His results can be interpreted in a way that is similar to the reasoning offered by some of my participants after choosing the empirically dominated "27": both sets of participants seem to try to "play it safe" (in the case of Greiner's study, safeguarding against a strongly dominated move by the other player that they deem to be extremely unlikely, in my case against a "selfish majority" of others). De Kwaadsteniet et al. (2023) show that irrelevant aspects can make participants look for heuristic asymmetric solutions to the problem they are facing. In my study, this is different: people have to coordinate on asymmetric solutions without any help in breaking the symmetry. Within the literature on behavioural rules, closely related papers are Crosetto et al. (2020), Wolff (2021), and Sontuoso and Bhatia (2021).

Crosetto et al. (2020) have their participants bid for an object with known value in a two-player auction against a computer player whose bid is determined by uniform randomization. After that, participants have to indicate the probabilities with which the computer's bid falls into one of five evenly-sized bins of possible bids. Even though the computer's uniform-randomization strategy is known to participants, a majority report "beliefs that have a peak in the interior of the range".

My study adds to the findings of Crosetto et al. (2020) in several ways. I distinguish 'centred-belief players' from other participants whose beliefs have a peak in the interior but are strongly asymmetric. I further document some limits to the phenomenon, as it does not seem to be present when actions do not appear in their natural order. Most importantly, I relate participants' beliefs to their actions (Crosetto et al. report only that actions and beliefs are correlated between participants). In particular, I show that 'centred-belief players' differ qualitatively in their behaviour from the rest of the population, that this differing behaviour affects the overall data pattern in two different games, and that the types' strategies in the two games are correlated within-participants.

Wolff (2021) studies symmetric pure discoordination games in which participants' pure strategies can be distinguished only by their labels and positions. The study finds that participants play the pure discoordination games as if they were giving up on reasoning strategically and betting on one of the strategies. Note that—unsurprisingly—this explanation fails to account for the data presented in this paper (Wolff, 2021, proposes that agents resort to betting-like behaviour only in case strategic thinking leads to the conclusion that any option is as good as any other, which will not be the case here; if participants have to bet on "27 €", "30€", or "33€" to receive a prize of 10 Euros in case a random draw selects the same option, 18% bet on 27, 38% bet on 30, and 44% bet on 33).

Similarly, Sontuoso and Bhatia (2021) study coordination games and hide-and-seek games in which options are denoted by natural-language words. They show that players often choose actions with a frequently-used word as a label when they have an incentive to match their opponent's action. In contrast, players who have to mismatch their opponent's action rely less on actions that have a frequently-used word as a label (see, *e.g.*, Mehta et al., 1994, Bardsley et al., 2010, Faillo et al., 2017, or van Elten & Penczynski, 2020, for studies on strategic reasoning in coordination games). I explicitly vary the salience of the options between treatments, without any apparent effects on the resulting behaviour. In that sense, the monetary incentives seem strong enough to dominate any 'prominent-number effects'.

A third literature strand looks at the stability of behaviour across different games, as, for example, in Georganas et al. (2015) or Rubinstein (2016). Georganas et al. (2015) show that participants exhibit little stability in terms of their estimated level-k types across two families of games (see Cooper et al., 2024, and Hyndman et al., 2022, for similar results). In contrast, Rubinstein (2016) finds positive correlations between a whole range of games in terms of 'decision styles': the fraction of a participant's "contemplative" choices in collections of nine games in most cases is predictive of the same participant's choice in a tenth game. Similarly, my findings also suggest that decision strategies may not be that different between games. Identifying 'centred-belief players' in a setting where players want to coordinate on different actions is predictive for choices in a setting where players benefit hugely from outsmarting the other player (and where there always is at least one player who would want to deviate from their action *ex-post*).

3. Experimental design and setups

3.1. Initial focus of the study

The Base treatment is the focal name-your-prize game with options "27", "30", and "33". The initial focus of the study was a comparison of Base with the game by Berger et al. (2016), which for this paper will act as a 'Placebo' treatment. In their game, many players each choose one of three amounts. The amount chosen by the lowest number of players is the "winning amount", and one player would be drawn to receive that amount for payment from the group of players who had chosen the amount. The data from this 'Placebo' treatment suggest that the basic setup and the participant pool are fine, as the data shows the same monotonically increasing pattern as the data from Berger et al. (2016, I even obtain the same proportion of "intermediate" choices—34.7% *versus* 34%—even though payoffs and payoff differences are very different: in the original game, they were 100€, 150€, or 200€).

3.2. Robustness checks

Because the project started as an exploratory exercise, the initial treatments were added to unrelated experiments. Having seen the surprising data pattern in the baseline treatment, I explore the finding's robustness in a variety of treatments. To keep things comparable, I kept running the treatments as an extra task in unrelated experiments for most treatments. Given this procedure, I did not run any power analysis but ran two to three sessions for most treatments, sometimes four, depending on the availability of unrelated experiments. Given the data from the first treatments, two to three sessions in most cases seemed to give a rather clear indication of whether the aggregate pattern would end up being non-monotonic or monotonic, which is the indication of robustness I was looking for. To look at the pattern's robustness, I run five types of Rob(ustness)-treatments:

- The Rob-Num treatments change the payoffs to have an uneven number at the middle (Rob-Num 1: 24–27–30) or to exhibit larger differences (Rob-Num 2: 24–30–36, and Rob-Num 3: 20–30–30).
- The Rob-Pay treatments are the first treatments in which all players are paid (payoffs being reduced to 5–6–7, 7–7.5–8, and 9–10–11 in Rob-Pay 1, 2, and 3, respectively).
- The Rob-Pay-$|\mathcal{S}|$ treatments, in addition, vary the number of alternatives (4, 5, and 5 in Rob-Pay-$|\mathcal{S}|$ 1, 2, and 3, respectively), and
- The Rob-Pool treatments are run in a completely different participant pool (former students of game-theory courses), also varying (in this case, hypothetic) payments (5.4–6.3–7.2, 6000–7000-8000, and 5000–6000–7000–8000–9000 in Rob-Pool 1, 2, and 3). Finally,
- The Rob-Pop treatments Rob-Pop 1 and 2 have participants simultaneously play a whole population of other participants. Rob-Pop 1 uses an 'opponent frame', presenting the setup as a two-player game first, followed by the explanation that participants would play the two-player game against all others in the session simultaneously (being paid their average payoff). Rob-Pop 2 uses a 'population frame' that tells participants they will be paid their chosen amount for every other participant in the session who chose a different amount, divided by the number of participants in the session. These treatments were run as part of another project, Folli and Wolff (2022). That study used different belief-elicitation frames (the 'opponent frame' asking for the opponent's behaviour, the 'population frame' asking about all other participants' behaviour, and a 'random-other frame' asking about one other participant's behaviour who would not be the participant's opponent) to understand psychological processes at work when coming up with one's beliefs. The above treatments initially were run to show that the frames affect also choices in a game and not only beliefs (we did not run a 'random-other' treatment of the discoordination game, mainly because it is not clear what that treatment would look like). The results of those sessions were included in the working-paper version Bauer and Wolff (2018, last paragraph of Section 4), but not in the final version of Folli and Wolff (2022, D. Bauer changed his name to Folli).

Looking at treatments Rob-Pay 3, Rob-Pool 1, 2, and 3 (and Out-of-Game, explained below) also allows us to observe whether the results differ when the game is the first or even the only part of the experiment. In addition, the game was the second main part of the experiment in a Strategies treatment also described in the following (the first part was an unrelated one-shot public-good task played with a different opponent).

3.3. Mechanism treatments

In addition to the robustness tests, I ran two treatments aimed at understanding the predominant data pattern:

- Strategies uses 2-player teams as a means to incentivize the disclosure of participants' reasoning in a 40–44–48 treatment. The approach, pioneered by Burchardi and Penczynski (2014), rests on the idea of letting one team member suggest an action for the game and write a justifying message to the other member, who then makes the final decision for the whole team. To economize on research money and participants, it is not revealed who in the team is the suggesting player and who is the player making the final decision, so that both players suggest an action and write a message. Importantly, though, the communication along the actual path of play remains a one-way communication.

  This setup yields two types of choices, suggestions and implemented actions after participants have received the suggestions and messages of their fellow team member. In terms of choice data, I focus on the suggestions as they are uninfluenced by

Table 1
Overview of the data.

Treatment	Options	Option 1	Option 2	Option 3	Option 4	Option 5	N.obs.	Lab/ Online	Reason/Particularities	Payment
Base	27–30–33	34.5	**20.9**	44.5			110	lab	Baseline treatment	one pair
Strategies	40–44–48	28.9	**26.3**	44.7			76	online	discoordination in teams, suggestions	all (random part)
		(25.0)	(15.8)	(59.2)					(discoordination in teams, team decisions)	
Beliefs	27–30–33	24.1	**20.7**	55.2			58	lab	belief elicitation (+ replication)	one pair
Out–of–Game	17–20–23△,‡	24.2	**21.2**	54.5			99	online	predictive power of ‘centred–belief types’	all (random part)
Rob–Num 1	24–27–30	34.6	**15.4**	50			52	lab	not having “30” in the middle	one pair
Rob–Num 2	24–30–36	29.8	**23.8**	46.4			84	lab	increasing the payoff differences	one pair
Rob–Num 3	20–30–40	19.0	**17.9**	63.1			84	lab	increasing the payoff differences further	one pair
Rob–Pay 1	5–6–7	17.5	30	52.5			40	lab	paying all	all
Rob–Pay 2	7–7.5–8	21.2	**10.6**	68.2			66	lab	paying all	all
Rob–Pay 3	9–10–11△	17.6	29.4	52.9			85	online	benchmark for survey	all (random part)
Rob–Pay-$\lvert S\rvert$ 1	6.7–7.2–7.7–8.2	6.6	32.9	**13.2**	47.4		76	lab	increasing the number of alternatives	all
Rob–Pay-$\lvert S\rvert$ 2	5–5.5–6–6.5–7	2.1	12.7	17.6	21.8	45.8	142	lab	increasing the number of alternatives	all
Rob–Pay-$\lvert S\rvert$ 3	6.2–6.7–7.2–7.7–8.2	2.3	9.1	17	**15.9**	55.7	88	lab	...without having ‘round’ payments	all
Rob–Pool 1	5.4 – 6.3 – 7.2△	25.5	**25.3**	49.2			308	online	‘replication’ amongst GT students	no payment
Rob–Pool 2	6000–7000–8000△	27.5	30	42.5			209	online	...with large differences	no payment
Rob–Pool 3	5000–6000–7000–8000–9000△	8.1	16.1	**14.4**	25.5	36.5	236	online	large differences, 5 options	no payment
Rob–Pop 1	27–30–33	29.2	**14.2**	56.6			106	lab	discoordination with everybody, ‘opponent frame’	one pair
Rob–Pop 2	27–30–33	38	**24.1**	38			108	lab	discoordination with everybody, ‘population frame’	one pair
‘Placebo’	27–30–33	18.4	34.7	46.9			98	lab	‘replication‘ of the Berger et al. game	the winner

Non-monotonicities are marked in boldface. Rob stands for robustness, Num for numbers (the available prizes), Pay for paying all, $\lvert S\rvert$ for the number of available actions, Pool for the participant pool, and Pop for population. △The game was the first or only task of the experiment. ‡There was a conversion rate of 4 experimental tokens to 1 Euro.

others (focusing on implemented choices instead would boost the non-monotonicity of the observed choice pattern strongly). The treatment was preceded by a plain-vanilla public-good experiment, and only one of the two parts was randomly selected for payment. At the end of the session, payoffs were converted to Euros at a rate of 2 points per Euro, and there was a show-up fee of 5 Euros.
- Beliefs adds an additional part to the Base setup. After the game, I ask participants to report their estimate of the other participants’ choice probabilities. Players receive 2 Euros if their estimate does not differ from the true proportion by more than 2 percentage points for any of the options.

3.4. Exploring heuristic centred-belief play across games

The final treatment explores whether heuristic centred-belief play is persistent across games:

- Out-of-Game is analogous to Beliefs but adds the 11–20 game. Specifically, participants play a 17–20–23 treatment in the first part (in points, with a conversion rate of 4 points per Euro). Then, they go through the belief elicitation from Beliefs, and finally, they play the 11–20 game. One of the two games is selected randomly for payment. The final payoff consists of the payoff from the selected game and the belief elicitation, plus a fixed participation fee of 2.50 EUR.

3.5. Additional procedural details

Except for the three Rob-Pool treatments run on http://gametheory.tau.ac.il/, all sessions were run with the local participant pool of the LakeLab at the University of Konstanz. I used z-Tree (Fischbacher, 2007) for the experiments, and hroot (Bock et al., 2014) for recruitment. For the online sessions of the Rob-Pay 3, Out-of-Game, and Strategies treatments, I additionally had to use z-Tree-unleashed (Duch et al., 2020). Table 1 presents an overview of all experiments, their particularities, and the data.

4. Predictions

For the predictions, I use the symmetric Nash-equilibrium and the popular models from behavioural game theory. I focus on the symmetric equilibrium (for the focal game, 79/299, 101/299, 119/299) because there is no way of breaking the symmetry for the players (and thus, I discard asymmetric equilibria as implausible). Given that the standard (logit) quantal-response-equilibrium, cognitive-hierarchy, level-k, salience-theory, and impulse-balance-equilibrium models all predict monotonically-increasing patterns, too, I relegate their description including the specific predictions to Section B of the Online Appendix.

The only popular behavioural-game-theory model that can account for non-monotonic patterns in some cases is noisy introspection. In essence, noisy introspection combines the assumption that agents follow a logistic-choice function in their decisions (as in Quantal-Response Equilibrium) with the idea that there are levels of reasoning and a fixed level-0 that randomizes uniformly (Goeree & Holt, 2004). Each player of level-k then responds to a population of only level-($k-1$) players, using the above-mentioned logistic-choice function. Using the parameter estimates in Goeree et al. (2018) for prediction, the model predicts a non-monotonic pattern for five out of the 14 non-monotonic treatments (fitting the model to the data yields five non-monotonic treatments, too).

The intuition for why noisy introspection can account for non-monotonic patterns is the following: in the focal game, level-1 will choose "33" most often, because it responds to a uniformly-mixing level-0 by a logistic-choice function. However, (depending on the parameter of the logistic-choice function) it may choose "30" often enough to make level-2 respond by choosing mainly "27". Depending on the parameter of the choice function and the distribution of levels of reasoning, the above may be enough to generate an aggregate choice pattern that is non-monotonic overall. While the belief-clustering analysis in Table 3 of Section 5.3 only gives a rough approximation of the model, the mechanics will nicely show in the first three lines of Table 3.

5. Results

5.1. The non-monotonicity of the aggregate-data pattern

As can be seen from the first row of Table 1, the frequencies of "27"-, "30"-, and "33"-choices in Base are 34.5%, 20.9%, and 44.5%, respectively. This is not just a statistical outlier from an inherently monotonic data-generating process. To see this, I focus on all three-option games together (all treatments but Rob-Pay-$|S|$ 1, 2, and 3, plus Rob-Pool 3). Amongst them, I observe a U-shape in 11 out of 14 treatments. Abstracting from noisy introspection, the closest probability distribution to a U-shape that (at least) some of the models accommodate as a special case is uniform randomization. Under uniform randomization, the likelihood of observing at least 11 out of 14 U-shaped data sets given the sample sizes I used is 0.2%. In other words, I can reject even uniform randomization at the standard significance levels.

If the data cannot be explained by either noise or any of the behavioural models that predict monotonic choice patterns, can it be explained by noisy introspection (alone)? The answer is negative, too. To assess the question, I perform the following test: I fit the noisy-introspection model to each treatment individually and use the fitted probabilities as a data-generating process for 100,000 random data sets (so the null-hypothesis is that the data come from the fitted noisy-introspection model). Then I use the mean-squared deviation from the fitted probabilities as a statistic and compare the mean squared deviations of the actual data sets from the fitted probabilities to the distribution of mean squared deviations generated under the fitted noisy-introspection model. Only slightly more than 1% of the 100,000 generated data sets have a mean squared deviation that is at least as large as the actual data set. Imposing more structure (such as assuming a common Poisson parameter across treatments) reduces this simulated p-value even further. In other words, also the noisy introspection model cannot accommodate aggregate behaviour. Having looked at the general data pattern across all (three-option) treatments, let me explore the treatments in a more fine-grained way next.

Changing the prizes does not eliminate the non-monotonicity. The first check was to remove "30" from the centre (Rob-Num 1). Changing the options to 24–27–30 (slightly) increased the non-monotonicity rather than reducing it. Second, I increased the differences, from 27–30–33 in Base over 24–30–36 to 20–30–40 in Rob-Num 2 and 3, respectively. While there is little difference between Base and the Rob-Num 2 data, the non-monotonicity becomes weaker in the Rob-Num 3 treatment. However, it does not do so in the expected way: increasing the differences does not increase the prevalence of middle-option choices in the direction of the theoretic predictions. Instead, there is a shift away from the lowest option that in the most extreme Rob-Num 3 treatment goes entirely to the highest option.

The non-monotonicity does not depend critically on specifics of the payment rule nor on the size of the action space. Paying all (for sure) eliminates the non-monotonicity in one of the instances (Rob-Pay 1) but not the other (Rob-Pay 2). The same holds true for the 5-option treatments Rob-Pay-$|S|$ 2 and 3. Paying all in case the part is randomly drawn for payment also acts similarly: it eliminates the non-monotonicity in the Rob-Pay-3 case but does not do so in the Out-of-Game or Stategies treatments.

Introducing additional options leads to a similar picture as switching to paying all: In two out of three lab treatments, Rob-Pay-$|S|$ 1 and 3, the distribution exhibits a dip at the second-highest option, and in Rob-Pay-$|S|$ 2, the highest option still is chosen far too often compared to the standard equilibrium prediction.

The non-monotonicity is not specific to the pool of participants. Using a larger sample of former students of Ariel Rubinstein's game-theory courses moves the data closer to a monotonic pattern, too. However, in two out of three treatments, Rob-Pool 1 and 3, the data still exhibit a non-monotonicity in terms of a dip, even though in Rob-Pool 3, the dip occurs at the third-highest option.

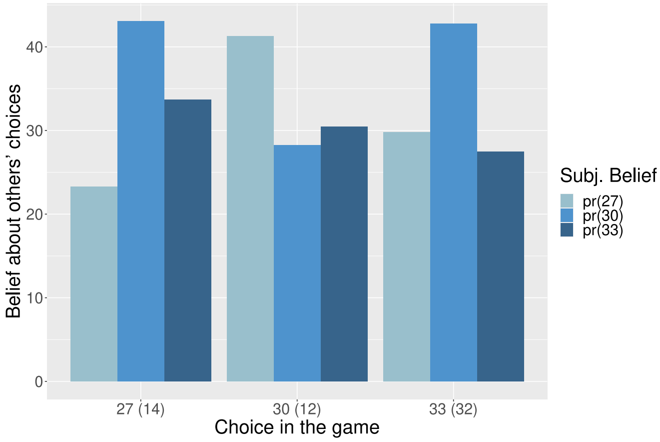

Fig. 1. Average beliefs, by participants' own choice (numbers of observations in parentheses).

Table 2
Participants' choice, by the best response to their reported belief.

Best response (N. obs.)	Actual choice, in (row-wise) percentages		
	"27"	"30"	"33"
"27" (9)	55.6	11.1	33.3
"30" (7)	0	57.1	42.9
"33" (42)	21.4	16.7	61.9

The non-monotonicity does not depend on the game being added as an extra part to other experiments. Finally, the non-monotonicity is not a consequence of conducting the treatments as supplementary parts to other experiments: amongst the five treatments that had the game as the first or only part of the experiment, again three led to a non-monotonicity.

Evaluation: The non-monotonicity result seems relatively robust to variations of the experiment. Summing up, none of the variations reliably eliminates the non-monotonicity. I observe a non-monotonicity in 14 out of 18 treatments, and compared to the symmetric Nash equilibrium, the second-highest option is 'under-played' in *all* 18 treatments. Similar findings hold for the other models, except for noisy introspection; noisy introspection is able to fit a non-monotonic pattern in 5 out of the 18 treatments. Taken together, these observations clearly speak in favour of taking the general phenomenon seriously. To understand behaviour in games, it is often very helpful to look at beliefs on top of actions. This is what the next two Sections are focused on.

5.2. Best-responses (based on participants' belief reports)

In Beliefs, I elicit participants' beliefs after they make their choice. They are paid an additional 2 Euros if their estimates of others' behaviour do not differ for any of the options by more than 2 percentage points (admittedly, this was a very difficult target, and only a single person's belief was so accurate). Fig. 1 shows the average beliefs, depending on participants' choices. Contrary to the actual choice pattern, the majority of participants seem to agree that most others will choose "30" ("30" carries the largest probability mass in the average belief of the '27–choosers' as well as that of the '33-choosers'). If we now look at what participants should have chosen given their belief, we obtain Table 2. As we can see on the diagonal, 56%–62% of the participants choose a best-response to their belief in terms of expected payoffs. While there is a slight increase from the first row to the last, the best-response rate is similar across rows.

Looking at choices that were not best-responses, roughly 30% play their second-best choice (43% for those whose best-response is "30"), and 10% play their 'worst-response' (the figures cannot be read from Table 2). These figures suggest that, by and large, participants are acting in agreement with their beliefs, subjects to decision-errors of a quantal-response/logit-choice type. This suggests that the reason for QRE's failure to predict the data lies in the equilibrium assumption rather than in the way beliefs translate into actions. And conversely, quantal-response behaviour seems to be an integral part of the explanation for participants' choices. So far, we have looked at individual belief reports. The next session looks at whether there are special types of participants in terms of their beliefs, to see whether I can relate the belief types to the types from any of the game theoretic models I considered in Section 4.

5.3. Heuristic centred-belief players

A cluster analysis of participants' beliefs reported in Table 3 and visualized in Fig. 2 yields additional insights. First, adding up the first three lines (for NI-1, NI-2, and NI-3) suggests that slightly more than half of the participants' beliefs correspond to a type

Table 3
Results of a cluster analysis of participants' beliefs, alongside the corresponding choice frequencies (best-responses are underlined). The labels in the right-hand column correspond to those of Fig. 2.

Belief on...			Frequency	Actual choice			Classification
"27"	"30"	"33"	(in %)	"27"	"30"	"33"	
32.0	35.7	32.3	34.5	4	3	13	≈ NI-1
18.4	33.7	48.0	13.8	5	2	1	NI-2
41.5	14.8	43.7	5.2	0	2	1	NI-3
22.0	52.0	26.0	22.4	5	1	7	CB ('centred belief')
36.7	46.1	17.2	12.1	0	1	6	Q1 (interpretation?)
52.2	24.7	23.1	6.9	0	2	2	Q2 (interpretation?)
7.5	92.5	0	3.4	0	0	2	out1 (outlier)
98	0	2	1.7	0	1	0	out2 (outlier)

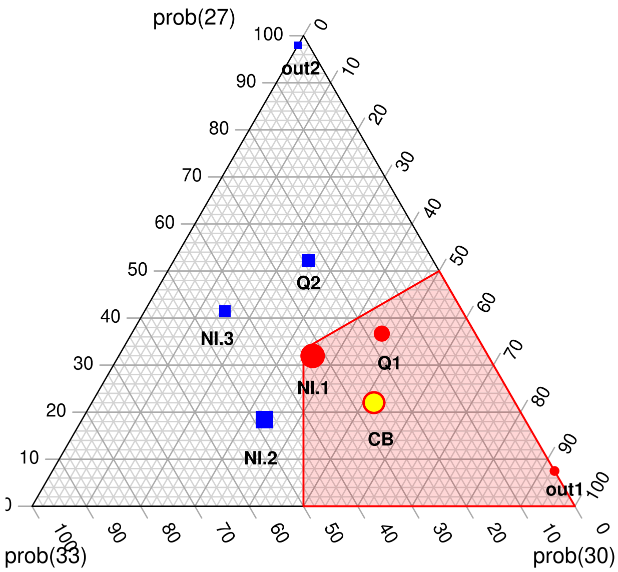

Fig. 2. Graphical representation of the belief clusters from Table 3. The red/shaded polygon marks all beliefs that have the largest probability mass on "30", and the 'centred-belief' (CB) type is marked by its yellow/light filling. There are three clusters that are roughly symmetric: NI-3, NI-1, and CB. However, amongst those three clusters, only the CB cluster has a clear mode on the probability mass assigned to "30". On the other hand, while the clusters denoted by "Q1" and "out1" also have the mode on "30", they are asymmetric.

from the noisy-introspection model. Second, looking at lines 4, 5, and 7, just short of 40% believe that "30" is the choice most others will choose (and about three quarters if we take the exact figures in row 1 by their face value; note that the number of clusters was determined using the SD index, which corresponds to a weighted sum of the average "point scattering" within clusters and the inverse of the total separation between clusters).

More importantly, though, almost a quarter of the population can be classified as having a 'centred belief' - a belief that places the highest probability on the central option and is (roughly) symmetric around it. And looking at the distribution of actual choices in columns 5–7, it is the 'centred-belief players' who are driving the non-monotonicity of the overall pattern (the only other groups exhibiting a slight dip are the NI-1s, as well as the combination of NI-types). Note that the identification of the 'centred-belief' cluster does not depend on the usage of the SD index: the 'centred-belief' cluster is the only cluster whose identification is robust to using different numbers of clusters (tested for a single cluster all the way to 10 clusters). In addition, the robustness check reported in Table A.1 in the Online Appendix suggests that the 22.4% shown in Table 3 are a *lower* bound for the 'centred-belief' cluster's prevalence.

5.4. Heuristic centred-belief players in other data sets

To test how relevant the result of a meaningful population of centred-belief players is, I looked for another data set of a game with a limited number of options that would have a natural order. The very obvious candidate was Arad and Rubinstein's (2012) 11–20 game in which each player would receive the amount they request, plus a bonus of 20 in case the other player chooses exactly one monetary unit more. Goeree et al. (2018) provide a data set that includes elicited beliefs for the third game that participants play. In the following, I run a cluster analysis on their data.

Table 4
Choice percentages in the standard 11–20 game by belief cluster (Goeree et al. (2018) elicited beliefs only for the third round participants played, and so, I also display the choice percentages from that round).

	"11"	"13"	"15"	"16"	"17"	"18"	"19"	"20"
Centred-belief players ($n = 11$)	0	0	36	27	9	9	9	9
Other types ($n = 61$)	2	2	3	2	23	38	23	8

Table 5
Choice percentages in the standard 11–20 game, by belief cluster from the 17–20–23 treatment.

	"13"	"14"	"15"	"16"	"17"	"18"	"19"	"20"
Centred-belief players ($n = 25$)	4	8	0	16	4	52	12	4
Other types ($n = 73$)	0	1	1	7	19	47	16	8

Running a cluster analysis on Goeree et al.'s (2018) data from the standard version of the 11–20 game (determining the number of clusters again by the SD index) yields 4 clusters, including a 'centred-belief' one (this cluster is slightly asymmetric but is clearly hump-shaped, with average beliefs of 2%, 0%, 4%, 8%, 13%, 19%, 22%, 18%, 8%, and 5% for choices of 11 through 20, respectively). The cluster comprises 15% of the population. Notably, this cluster appears only in the variant in which the options follow their natural order, but not in a treatment in which the order is strongly perturbed (19–18–17–…–12–11–20). This means that options being naturally ordered seems to be a pre-condition for participants to come up with 'centred beliefs'.

Again, the centred-belief players exhibit a rather peculiar choice pattern, as can be seen from comparing the first row of Table 4 with the row representing all other types. Strikingly, the majority of choices from centred-belief players are below 17 (representing 58% of all such choices; recall that they make up for only 15% of the population; the qualitative finding is quite robust: using, for example, 6 (10) clusters, the numbers would change to 17% (8%) of the population making up for 2/3 (1/2) of the choices below 17).

In contrast, only 8% of the choices of other players are below 17 (a Boschloo test on whether choices below 17 are equally prevalent amongst centred-belief players and other types yields $p < 0.001$; note that the Boschloo test is a uniformly more powerful alternative to Fisher's exact test). Focusing on choices only from the first round of play in their setup, the 15% centred-belief players still account for 36% of the choices below 17. Note that these choices are clearly dominated: judging by the (clustered) belief, the optimal choice of 17 yields a payoff (20.8) that surpasses the most common choice in that group, 15 (payoff: 18.9), by almost 2, whereas the numbers 18 and above all yield payoffs of at least 19.5. This means that in the 11–20 game, centred-belief players clearly do not follow the general quantal-response pattern.

Having found the centred-belief type in two different games, I set out to test the robustness of the findings in a new experiment, and whether classifying participants as centred-belief players in one game is predictive of their choices in the other. The following Section shows that this indeed is the case.

5.5. *Centred-belief play as a general strategy*

To analyse the out-of-game predictive power of the centred-belief-player categorization, I ran the Out-of-Game treatment that had two parts. The first part is a 17–20–23 treatment with belief-elicitation, and the second part is the 11–20 game. The hypothesis to be tested was that participants classified as centred-belief players by their beliefs in the first game would again make up for a disproportionally large share of the choices below 17 in the 11–20 game. Recall that in Goeree et al.'s data, '11–20-game centred-belief players' made up for 15% of the population but for 58% of the choices below 17.

Table 5 presents the results for the across-game analysis (I had to drop one out of 99 participants because, due to a programming error, the participant managed to get past the screen of the 11–20 game without making a choice). While the effect is weaker, it clearly remains present: centred-belief players are 3 times as likely to choose a number below 17 compared to all other participants (28% as opposed to 9.6%), again making up for 50% of all such choices (a Boschloo test on whether choices below 17 are equally prevalent among centred-belief players and others yields $p = 0.03$).

The findings show two things. First, I again find a substantial fraction of centred-belief players (that is very similar in size to what I observed among the 58 earlier participants) who produce a non-monotonic pattern, choosing "17", "20", and "23" with relative frequencies of 32%, 12%, and 56%, respectively. And second, in the sense of the tested hypothesis, reporting a 'centred belief' in one game has predictive power for behaviour in another.

5.6. *Another glimpse into players' minds: team communication*

Before running the cluster analyses on participants' beliefs reported in the preceding Section, I ran three additional sessions of the Strategies treatment aimed at incentivizing reports of strategic reasoning. While the results are not very informative about the distinction between centred-belief players and other types, they provide an additional perspective on participants' reasoning. I therefore include a brief description of the data in the following final results section.

Following Burchardi and Penczynski (2014), in Strategies the game was played in pairs. Each pair has a suggesting player and a deciding player. The suggesting player suggests an action in the game and writes a free-form message that is transferred to the deciding player alongside the suggested action. The deciding player sees both the suggestion and the message before deciding on the team's action, which gives the suggesting player an incentive to give a compelling argument for why the suggested action is a good choice (recent evidence by Cavalcanti et al., 2023, suggests that leaders' suggestions are followed particularly often when they have information that followers do not have, even if this information is irrelevant).

Also following Burchardi and Penczynski (2014), I use the strategy method with respect to the player roles to save on participants (*i.e.*, both members of the team might be assigned either role. Thus, both have to supply a suggestion-message pair and make a final decision after seeing their partner's message; importantly, along the path of action, communication remains one-way). The possible actions were 40, 44, and 48, which were to be split between the two team members in case of success.

As reported in Table 1, 29%, 26%, and 45% suggested 40, 44, and 48, respectively (the team decisions were 40, 44, and 48 with probabilities of 25%, 16%, and 59%). More importantly, though, 41% of those who suggest 40 explicitly mention "safety" in their messages. This consideration is virtually absent in others' explanations: only 4% of those who suggest 44 or 48 refer to "safety" (and another 4% to being "careful"). In contrast, 19% refer to "going for the risk" (although in more varied terms including "All in" or "No risk no fun"), which is absent in the former group's messages.

The notion of the lowest option being "safe" clearly suggests a belief that everybody else will be choosing either 44 or 48. This belief would correspond to a level-2 player in Level-k based on a level-0 that chooses by the option suggesting the highest payoff (Arad and Rubinstein's, 2012, argument for using "20" as the level-0 choice in their 11–20 game), or an NI-2 player in the noisy-introspection model. Note, however, that this observation does not square up with the numbers from the cluster analysis on reported beliefs: According to Table 3, clearly less than 20% belong to a belief cluster that would suggest "27" to be a "safe" choice—which, empirically, it is not.

6. Conclusion

In this paper, I introduce the name-your-prize game, in which two players choose a monetary amount from a given set and obtain the amount if and only if their opponent chooses a different amount. Studying this game is interesting for two reasons: (i) it yields an aggregate choice pattern that none of the prominent models of behavioural game theory predicts consistently; and (ii) it allows us to identify 'heuristic centred-belief players' as a new type of player in the population whose behaviour explains the above pattern. As this paper shows, the type should not be neglected, either in terms of its prevalence or in terms of its effect on the other players' empirical best-response in the game (which is a consequence of the type's effect on the aggregate choice pattern).

I call the observed type of players 'centred-belief players' because they have a roughly-symmetric belief that peaks in the centre of the option set. The fact that some participants display a 'centred belief' is not surprising: it simply is a specific form of a central-tendency bias (note that in the cluster analysis in Section 5.3, there is an additional belief cluster that also peaks in the centre, but that is highly asymmetric, denoted as "Q1" in Table 3 and Fig. 2). 'Centred beliefs' may come from a pre-conception that 'most variables lead to a Gauss curve' (think of age, IQ scores, …), which would likely be triggered only if the underlying variable has a natural ordering.

I call the type 'heuristic' because in two out of three experiments (Beliefs and Goeree et al.'s, 2018, 11–20 data, but not Out-of-Game), their propensity to play a best-response to their stated beliefs is clearly lower than the propensity of other types. As a consequence, best- or quantal-responding to their beliefs cannot yield a satisfactory explanation for the type's behaviour (which sets the type apart from, *e.g.,* level-k or noisy-introspection types). In that, the players' response to their belief may be driven by a reasoning of 'playing it safe' which, however, leads to making an empirically dominated choice.

Seeing the 'heuristic centred-belief type' in two rather different games suggests that it is a general phenomenon to some degree. Seeing the predictive power of players' 'centred beliefs' in one game for their actions in another game shows that a certain fraction of the population uses 'heuristic centred-belief play' as a general strategy within a certain type of games.

The findings are important because they (a) help to understand the behaviour of a non-negligible share of the population, (b) because we need to take the player type into account when making predictions for new settings, and (c) because the findings show that we cannot look only at aggregate data when searching for the best model of behaviour. They show that we even cannot look only at individual choices: Arad and Rubinstein (2012) claim that the 11–20 game transparently and unambiguously identifies participants' levels of reasoning because "[i]t is hard to think of plausible alternative decision rules for this game. (...) The only other conceivable rules of behaviour we could think of are randomly choosing a strategy or arbitrarily guessing the other player's strategy and best-responding to it." As I show in this paper, 'heuristic centred-belief players' neither choose their strategy "randomly" (at least not in the usual sense) nor do they best-respond to arbitrary guesses, and yet, they do follow an alternative decision rule. What alternative decision rule they follow, and exactly which features of a situation trigger 'heuristic centred-belief play' is an important topic of further research.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.joep.2025.102806.

References

Achtziger, A., & Alós-Ferrer, C. (2014). Fast or rational? A response-times study of Bayesian updating. *Management Science*, *60*(4), 923–938.

Alós-Ferrer, C., & Buckenmaier, J. (2021). Cognitive sophistication and deliberation times. *Experimental Economics*, *24*(2), 558–592.

Arad, A., & Rubinstein, A. (2012). The 11–20 money request game: A level-k reasoning study. *American Economic Review*, *102*(7), 3561–3573.

Bardsley, N., Mehta, J., Starmer, C., & Sugden, R. (2010). Explaining focal points: Cognitive hierarchy theory *versus* team reasoning. *Economic Journal*, *120*, 40–79.

Bauer, D., & Wolff, I. (2018). Biases in beliefs: Experimental evidence. TWI Research Paper Series, No. 109.

Berger, U., De Silva, H., & Fellner-Röhling, G. (2016). Cognitive hierarchies in the minimizer game. *Journal of Economic Behavior and Organization*, *130*, 337–348.

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, *71*, 117–120.

Burchardi, K. B., & Penczynski, S. P. (2014). Out of your mind: Estimating individual reasoning in one shot games. *Games and Economic Behavior*, *84*, 39–57.

Cavalcanti, C., Grossman, P. J., & Khalil, E. L. (2023). Leadership heuristic. *Journal of Economic Psychology*, *98*, Article 102661.

Cooper, D. J., Fatás, E., Morales, A. J., & Qi, S. (2024). Consistent depth of reasoning in level-k models. *American Economic Journal: Microeconomics*, *16*(4), 40–76.

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

Crosetto, P., Filippin, A., Katuščák, P., & Smith, J. (2020). Central tendency bias in belief elicitation. *Journal of Economic Psychology*, *78*, Article 102273.

De Kwaadsteniet, E. W., Gross, J., & van Dijk, E. (2023). A "More-is-Better" heuristic in anticommons dilemmas: Psychological insights from a new anticommons bargaining game. *Journal of Economic Psychology*, *98*, Article 102653.

Duch, M. L., Grossmann, M. R., & Lauer, T. (2020). z-Tree unleashed: A novel client-integrating architecture for conducting z-Tree experiments over the internet. *Journal of Behavioral and Experimental Finance*, *28*, Article 100400.

Faillo, M., Smerilli, A., & Sugden, R. (2017). Bounded best-response and collective-optimality reasoning in coordination games. *Journal of Economic Behavior and Organization*, *140*, 317–335.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Folli, D., & Wolff, I. (2022). Biases in belief reports. *Journal of Economic Psychology*, *88*, Article 102458.

Georganas, S., Healy, P. J., & Weber, R. A. (2015). On the persistence of strategic sophistication. *Journal of Economic Theory*, *159*(A), 369–400.

Goeree, J. K., & Holt, C. A. (2004). A model of noisy introspection. *Games and Economic Behavior*, *46*(2), 365–382.

Goeree, J. K., Louis, P., & Zhang, J. (2018). Noisy introspection in the 11–20 game. *The Economic Journal*, *128*(611), 1509–1530.

Greiner, B. (2023). Strategic uncertainty aversion in bargaining — Experimental evidence. *Journal of Economic Psychology*, *95*, Article 102604.

Hyndman, K., Terracol, A., & Vaksmann, J. (2022). Beliefs and (in)stability in normal-form games. *Experimental Economics*, *25*, 1146–1172.

Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *American Economic Review*, *84*(3), 658.

Rubinstein, A. (2016). A typology of players: Between instinctive and contemplative. *The Quarterly Journal of Economics*, *131*(2), 859–890.

Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organization*. Macmillan.

Sontuoso, A., & Bhatia, S. (2021). A notion of prominence for games with natural-language labels. *Quantitative Economics*, *12*(1), 283–312.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

van Elten, J., & Penczynski, S. P. (2020). Coordination games with asymmetric payoffs: An experimental study with intra-group communication. *Journal of Economic Behavior and Organization*, *169*, 158–188.

Wolff, I. (2021). The lottery player's fallacy: Why labels predict strategic choices. *Journal of Economic Behavior and Organization*, *184*, 16–29.