LEMMA Working Paper

n° 2025-02

Reasonless Choices in One-Shot Games

José De Sousa

Université Paris-Panthéon-Assas, LEMMA

Guillaume Hollard

Ecole Polytechnique, CNRS

Antoine Terracol

LED, Université Paris 8

LEMMA EA 4442 - 4 rue Blaise Desgoffe 75006 Paris

Reasonless Choices in One-Shot Games

José De Sousa, Guillaume Hollard and Antoine Terracol*

May 8, 2025

Abstract

Experimentalists frequently encounter decisions that defy rational explanation. However, the reasons why subjects fail to make sensible choices remain debated. We propose a lab-in-the-field experiment that allows us to detect "reasonless choices" and evaluate several explanations for their occurrence. We find that most common explanations, such as low incentives and cognitive ability, have limited explanatory power. Our results underscore the need for deeper investigation into why some individuals fail to make sensible decisions.

*José De Sousa: Université Paris Panthéon-Assas, LEMMA and Sciences Po, LIEPP, France. Email: jose.de-sousa@u-paris2.fr. Guillaume Hollard: Ecole Polytechnique and CNRS, guillaume.hollard@polytechnique.edu. Antoine Terracol: LED, Université Paris 8, terracol@univ-paris8.fr.
We are especially grateful to Gary Charness for his encouragement to revise the paper and for pushing us to better understand why participants make choices that are hard to rationalize. His thoughtful suggestions, many of which we have incorporated, have shaped the paper into what it is today. We also thank G. Aldashev, J. Andreoni, G. Attanasi, A. Caplin, O. Compte, G. Coricelli, M.P. Dargnies, G. Frechette, T. Gajdos, U. Gneezy, H. Maafi, G. Mayraz, M. Niederle, T. Offerman, S. Sadoff, J. Sobel, L. Santos Pinto, T. Verdier and participants at ASFEE, Berlin Schumpeter Seminar, Copenhagen U., Ecole Polytechnique, EEA, MBB on-going research seminar, NYU CESS seminar, Paris-Saclay U., Paris School of Economics TOM seminar, PEJ, Rady UC San Diego seminar, Rennes U., and SBDM for comments. A. Fischman and E. Kaddouch provided excellent research assistance. N. Bonzou and A. Clauzel from the French Chess Federation provided decisive help.

There is nothing wrong with this paper – Gary Charness[1]

1 Introduction

A reasonless choice is a decision that an external observer struggles to explain because it defies any coherent rationale associated with deliberate decision-making. The prevalence and causes of reasonless choices remain debated. Why do some decision makers fail to make deliberate decisions? Experimental economics, especially through simple one-shot games, provides an opportunity to get insights on the origins of reasonless choices. In particular, we propose to test common explanations for reasonless choices, including unobservable information, limited cognitive ability, insufficient incentives, lack of attention and lack of expertise.

Our contribution is twofold. First, we introduce a novel identification strategy that establishes a lower bound on the fraction of reasonless choices by proposing a minimal consistency principle that any deliberate strategy should satisfy. Second, our lab-in-the-field, or extra-laboratory experiment as defined by Alekseev et al. (2017), shows that the substantial fraction of reasonless choices we observe cannot be accounted for by the most common explanations.

We focus one-shot games for two main reasons. First, they provide a clean environment to observe decision-making without the possibility of learning from feedback or repeated interaction. Because players have no chance to adjust their strategies over time, any confusion or random behavior is more easily detected. One-shot games are especially useful for studying the origins and prevalence of reasonless choices. Second, we hypothesize that experts, who bring specialized skills or strategic ability, might navigate one-shot games more effectively. In that respect, we recruited chess players who are often thought of as experts in strategic reasoning.

[1]This paper owes much to the intense discussions we had with Gary Charness, whose support and suggestions greatly improved its development. Gary initially raised several objections to an earlier version of the manuscript. Sitting together in a bar in Paris, we boldly claimed that we could address all his concerns. After a few drinks and two hours of spirited debate, Gary (reluctantly) conceded, "There is nothing wrong with this paper." Even more encouraging, he offered to monitor its progress and invited one of the co-authors to give a talk at the University of Santa Barbara, to foster further discussion.

Our paper addresses two sets of questions:

1. *How frequent are reasonless choices, and how can they be detected?* Classifying a choice as reasonless is challenging because the observed outcome alone may not reveal whether the decision was made deliberately or randomly. Consider a beauty contest game in which participants choose numbers between 0 and 100, aiming to guess the number closest to a fraction (typically 2/3 or 4/3) of the group average. Some subjects may choose 25 as a best-response to their beliefs about others' behavior. However, other subjects may also pick 25 because it is their lucky number or because the game seems too complex, prompting them to choose randomly.

   Our experiment offers an intuitive method to identify reasonless choices by eliciting behavior across parametric variations in a series of games. Subjects who apply a deliberate rule (e.g., equilibrium choices, heuristics, etc.) are expected to exhibit minimal consistency across games. In contrast, subjects who fail to demonstrate such consistency can hardly be considered as making a deliberate choice and their choices are classified as "reasonless". For example, one subject might choose 25 in every beauty-contest game regardless of whether the target is set at 2/3 or 4/3 of the average, while another subject might choose 25 with 2/3 and 75 with 4/3, reflecting deliberate strategic behavior. Comparing behavior across games thus allows us to distinguish between choices made deliberately and those that are essentially random. Using our approach, we find that 30% of choices can be classified as reasonless.

2. *What drives the prevalence of reasonless strategies, and how can it be reduced?* We find that common explanations in the literature, such as low incentives (Yechiam & Zeif 2023), poor instructions, and low attention, do not appear to drive the high frequency of reasonless choices. In our experiment, subjects show attention, understand instructions, and care about incentives. Furthermore, using chess players ranging from beginners to grandmasters, we find that strategic ability is only weakly correlated with the use of reasonless strategies. Even very strategy savvy subjects, such as chess grandmasters, often resort to reasonless choices in simple one-shot games. Predicting the quality of decisions based on observable characteristics thus

remains a critical challenge.

The remainder of the paper is organized as follows. Section 2 reviews the literature and outlines methods that demonstrate the prevalence and significance of reasonless choices in one-shot games—games. These findings are robust across experiments, subject pools, and game types. Section 3 describes our experimental design, which comprises three phases. Section 4 presents our findings, and Section 5 examines the most common explanations for reasonless choices. Finally, Section 6 concludes.

2 Review: Detecting Reasonless Choices

The simplest method to identify reasonless choices is to examine dominated strategies, as they cannot be best responses to any beliefs. However, relying solely on dominated strategies may overlook many cases. For example, in games without dominated strategies, reasonless choices would go undetected. Different methods have been designed to detect choices that do not result from an attempt to form beliefs and best-respond to them. The challenge, as noted above, is to distinguish between identical choices made for different reasons. We describe six empirical strategies designed to differentiate reasonless choices from deliberate ones. These methods show that between 30% and 50% of choices are reasonless.

1. Costa-Gomes & Crawford (2006) examine the consistency between stated beliefs and behavior. In a series of 14 two-person games, subjects report their beliefs and chose a strategy. Ideally, the chosen strategy should be the best response to these beliefs; however, more than half of the subjects fail to best respond to their own stated beliefs in at least half of the games they play.

2. In Agranov et al. (2012), subjects play beauty contest games against human opponents and computers. The computers select random strategies using a uniform distribution. Therefore, increasing the proportion of human opponents (and reducing the number of computers) should prompt a change in chosen strategies. However,

half of the subjects fail to best respond to computers, and about 30% of subjects do not revise their strategies when the composition of the player pool changes.

3. Burchardi & Penczynski (2014) implement a protocol in which subjects exchange messages and give advice on how to play. They classify the messages according to their strategic content (e.g., whether they refer to a best response). They find that one-third of the players send messages that do not include any mention of best-response reasoning.

4. In Ivanov et al. (2010), subjects play against their past-selves. In the second phase of the game, they are instructed that they are facing their phase one strategies. This design identifies which players can recall best respond to their past strategy. Since strategies arising from a clearly defined cognitive process are more likely to be remembered, this design enables classification of strategies. About one-third of the subjects are unable to simply replicate their past action.

5. Agranov et al. (2015) design an experiment in which players first make an immediate choice and then have the opportunity to revise it after additional time to think. A single point in time is randomly selected for payment, ensuring that players have incentives to report their best strategy at every stage. Their design tracks the decision-making process over time in a beauty-contest game. 45% of subjects persist in using dominated strategies even after ample reflection.

6. Enke et al. (2024) elicit a measure of cognitive uncertainty for each subject. They observe a robust pattern across 30 choice problems: high cognitive uncertainty is associated with high attenuation, i.e., subjects' decisions show low responsiveness to the fundamentals of the choice problem. Since a substantial fraction of subjects express high cognitive uncertainty, reasonless choices are likely to occur in a large set of choices problems. Moreover, since the 30 problems were selected as particularly representative of economic decisions, these findings suggest that reasonless choices are not confined to strategic interactions.

We now introduce a novel identification strategy that proposes a minimal consistency principle that any deliberate strategy should satisfy.

3 Experimental Design

3.1 Experts in Strategy?

Experts are individuals with specific abilities, often due to their roles in society or the economy. The behavior of top experts is particularly interesting because it sometimes diverges from that of students. For instance, experienced dealers are not prone to the endowment effect compared to students (See Engelmann & Hollard 2010). Since one-shot games require strategic ability, top chess players are often considered ideal candidates for bringing strategic behavior into the lab. Chess inherently demands strategic thinking, such as anticipating others' reactions to one's decisions.

The importance of "rationality spillovers" from chess expertise to one-shot games is a matter of debate.[2] Our aim is not to settle this debate but rather to highlight that if grandmasters struggle to behave rationally, it casts doubts on whether other populations would consistently do so in the lab. Thus, chess experts are are not chosen for their societal status but to test whether any group can meet the demands of non-cooperative game theory. In this context, using chess players serves more as a "nonexistence proof" than an exercise in external validity targeting a specific population.

We recruited 270 chess-players during a major international tournament held in Paris in 2010. Subjects were approached while not playing. They briefly interacted with the recruiter to confirm their proficiency in French (or English, for a very small minority). Participants were then directed to an adjacent room set up as an experimental lab. The experiment was computerized, with instructions displayed on the screen and also read aloud by the experimenter. Subjects were allowed to ask questions.

[2]See Levitt et al. (2010) and Palacios-Huerta & Volij (2009), and evidence from a beauty-contest game in Buhren & Frank (2012).

3.2 Phases of the Experiment

Our experiment consists of three phases:

Phase 1. Subjects play a series of 10 beauty-contest games. They choose a number as close as possible to $m \times$mean (where mean denotes the average of all players' answer). The parameter m takes on two values: $m = 2/3$ and $m = 4/3$.

Each game is played against five types of opponents, labeled as A, B, C, D and Random, where letters indicate the Elo ranking of chess players.[3] "Random" means the subject faces a device that selects strategies uniformly across the strategy space. Subjects play 10 games, one against each opponent type for each $m \in \{2/3, 4/3\}$, in random order, with no feedback provided during this phase.

All treatments were identical except that half of the subjects play against one opponent only, while the other half play against *two* opponents of the same level (A, B, C, D or Random). This distinction matters because the two-player version of the game has a dominant strategy, whereas the three-player version does not. In addition, the payment rule (10 points per game, shared among winners) creates a difference in expected earnings: subjects in the three-player version earn an average of 33.33 points, while those in the two-player version earn 50 points on average.

Phase 2: After playing ten beauty contests, subjects start a new game, the 11-20 game (described below), played once against another chess player of unspecified level. Before starting, subjects answer questions to confirm they understand the rules.

After completing all eleven games (the ten beauty contests and the 11-20 game), the screen displays the numbers chosen by the players in each game. Subjects can then observe the consequences of their actions, with each choice earning a certain number of points. These points are converted into Euros at the previously announced exchange rate of .2€ per point. Finally, subjects proceed in turn to another room to receive their cash payments anonymously.

Phase 3: *After* receiving their cash payment, subjects are offered the chance to participate in an additional beauty-contest game (with $m = 2/3$) involving all participants.

[3] A = Elo $>$ 2150, B = 2150 $>$ Elo $>$ 1800, C = 1800 $>$ Elo $>$ 1500, D = Elo $\leq$ 1500. The higher the Elo rating, the stronger the player.

Subjects are informed that the name of the winners will be publicly announced at the end of the 10-day chess tournament.[4] The two best players (i.e., those closest to the winning number) each received a cash payment of 150€. By this stage, subjects have received feedback on their performance in the first two phases.

3.3 Theoretical predictions

3.3.1 The beauty-contest game

The beauty-contest game, employed in Phases 1 and 3, is widely used in game theory to illustrate step reasoning (see Buhren et al. (2012) for a historical overview). Each player i selects a number x_i between 0 and 100. The objective is to choose the number closest to the target $m \times \frac{1}{n}\sum_{i=1}^{n} x_i$, where m is a parameter and n is the number of players. The player whose choice is nearest to the target wins a fixed prize, while all other players receive nothing.

For $m < 1$, the unique equilibrium is for all players to choose 0.[5] We also consider a version with $m > 1$; here, the focal equilibrium is for all players to choose 100.[6]

One interesting feature of the beauty contest is that in two-player games, a (weakly) dominant strategy exists: playing 0 when $m < 1$ and 100 when $m > 1$. However, in games with three or more players, no dominant strategy exists.

In the popular case where $m = 2/3$, the mean value chosen in the literature is about 35, which is far from the equilibrium prediction. Almost no subjects play the equilibrium strategy in one-shot games, including chess players.

3.3.2 The 11-20 game

The 11-20 money-request game, used in Phase 2, is presented as follows (Arad & Rubinstein 2012):

[4] Winners' names are announced immediately after the official chess tournament results, meaning winners may wait up to 10 days for payment.

[5] If the strategy space is restricted to integers, all players choosing 1 can also form an equilibrium. In the case of a tie, players can share the prize, or it can be distributed randomly, as in our experiment. If the entire prize is given to all tied players, additional equilibria may emerge.

[6] With three or more players, an unstable equilibrium exists in which all players choose 0. See Lopez (2001) for more details on the equilibrium set for integer games.

> "You and another student are playing a game in which each player requests an amount of money. The amount must be an integer between 11 and 20 shekels. Each player will receive the amount he requests. A player will receive an additional 20 shekels if he asks for exactly one shekel less than the other player. What amount of money would you request?"

To win a premium a player must choose *exactly* one step lower than the other player. Given the structure of the game, there is no Nash equilibrium in pure strategy. Assuming both players maximize expected gains, a unique symmetric mixed-strategy Nash equilibrium exists.[7] The distribution of the symmetric equilibrium puts a zero probability on strategies 11 to 14, probability 1/4 on strategies 15 and 16, and probabilities 4/20, 3/20, 2/20 and 1/20 on strategies 17, 18, 19 and 20, respectively. This equilibrium distribution is not intuitive and depends on the assumptions made about players' utility functions. Arad and Rubinstein show that even students trained in game theory do not behave as theory predicts. Nevertheless, their results provide a benchmark for the performance of subjects expected to be among the most strategic.

4 Results

We first present our method for identifying reasonless choices. Given that some players consistently make these choices, we then investigate whether our method offers a novel approach to detecting "random players", i.e., those who consistently behave randomly.

4.1 Identifying Reasonless Choices

The cleanest application of our method emerges in beauty contest games played against a random device, where beliefs are exogenously fixed. If a player selects a higher number when $m = 2/3$ than when $m = 4/3$ against the same random device, such behavior is difficult to explain. Therefore, any choice where a player does not select a higher number when $m = 4/3$ is classified as reasonless. We extend this method to games played against

[7] There are four additional asymmetric mixed strategy equilibria.

human opponents, where we assume that a strategic player will not choose a higher number when $m = 2/3$ than when $m = 4/3$ while facing the same opponent. In our study, 270 players make 10 choices each (a total of 2700 choices), and our methodology classifies 810 of these as reasonless, roughly 30% of all choices.

Our method for classifying reasonless choices avoids complex belief-elicitation procedures, which are challenging and may obscure whether observed inconsistency stem from the elicitation process or from inconsistent beliefs. Note also that relying solely on dominated strategies would fail to capture a substantial portion of reasonless choices.

4.2 From Reasonless Choices to Reasonless Individuals?

Previous findings indicate a substantial fraction of reasonless choices. Here, we explore whether some players consistently fail to elaborate a sounding strategy and instead resort to random choices. Our ultimate goal is to tear apart accidental mistakes from systematic erratic behavior to better understand the origins of reasonless choices. We rely on three types of evidence to assess whether some players act in a consistently random manner. First, based on the structure of the 10 games, we examine whether a subgroup of players shows no variation in choices when m varies. Next, we explore the plausibility of independent and identically distributed choices across games. Finally, we perform out-of-sample exercises to evaluate the stability of behavior across different games.

4.2.1 Mistakes vs Random Choices: Do Random Players Really Exist?

Our empirical strategy is to isolate a sub-group of players who make reasonless choices and evaluate the extent to which they consistently employ random strategies. If such a group exists, we interpret reasonless choices as stemming from random decision-making rather than mere mistakes.

Based on our identification strategy, we classify players according to the number of times they play higher with $m = \frac{2}{3}$ than with $m = \frac{4}{3}$ against the same opponent.[8] With

[8] For completeness, note that some players may hold extreme beliefs about their opponents in the three-player version that could rationalize choosing a higher number in the $m = \frac{2}{3}$ game than in the $m = \frac{4}{3}$ game. Although theoretically possible, such cases are highly unlikely.

Table 1: PRAI Classification of Players

Index Value	Number	Frequency
0	15	5.6
1	21	7.8
2	26	9.6
3	61	22.6
4	46	17.0
5	101	37.4
Total	270	100.0

Notes: The Pairwise Rationalizable Actions Index (PRAI) quantifies the consistency of players' choices across five pairs of games. An index of 0 indicates that a player consistently chooses a higher number with $m = \frac{2}{3}$ than with $m = \frac{4}{3}$; an index of 5 indicates no such inconsistency.

five pairs of games against the same opponents (levels A, B, C and D, and Random), we obtain five observations per player, which we use to calculate a "Pairwise Rationalizable Actions Index" (PRAI). Players are distributed into six levels: an index value of 0 indicates that subjects systematically choose a higher number when $m = \frac{2}{3}$ than when $m = \frac{4}{3}$, while an index of 5 indicates no such inconsistency. Table 1 shows the distribution of players according to this index.[9]

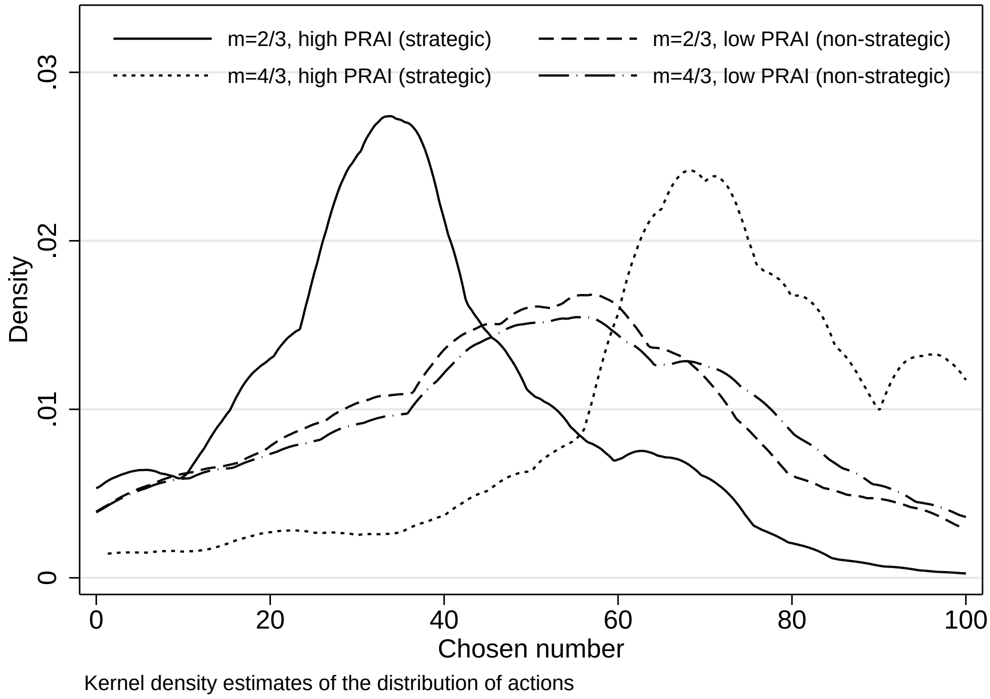
We divide players into two groups based on their PRAI: those with values from 0 to 3 (labeled "non-strategic") and those with values of 4 or 5 (labeled "strategic"). Figure 1 shows that non-strategic players exhibit little reaction to changes in m, averaging 48.98 for $m = \frac{2}{3}$ and 51.25 for $m = \frac{4}{3}$. In contrast, strategic players respond as expected, playing on average 36.12 when $m = \frac{2}{3}$ and 70.09 when $m = \frac{4}{3}$.[10] Fixed-effects regressions of choices on a dummy for $m = 4/3$, controlling for opponent type and period, yield an estimated coefficient of 2.19 (p-value=0.046) for non-strategic players and 33.91 (p-value10^{-3}) for strategic players. The difference in the explanatory power is also striking, with a within R-squared of 0.018 for non-strategic players and 0.528 for strategic players.

Next, Table 2 highlights the correlation of choices across games. If choices were com-

[9] Some readers might worry that our criterion, based solely on observed Phase 1 choices, could simply reflect how we split the sample rather than any true differences between players. In other words, it may be that all players are merely randomizing their actions, with some ending up with a high PRAI purely by chance. However, simulations presented in Appendix A show that if all players were random, the resulting PRAI distribution would differ markedly from our observed data.

[10] Descriptive statistics for strategic and non-strategic players are provided in Appendix Tables 7 and 8

Figure 1: Strategies Chosen by Low (Non-Strategic) and High (Strategic) PRAI Players



pletely random, they would be independent and identically distributed (i.i.d.) across games. The table presents the correlation coefficients between the five sequential choices for each value of m. For strategic subjects (top panel), the coefficients range from 0.649 to 0.791, whereas for non-strategic players (bottom panel) they are significantly lower, ranging from 0.153 to 0.372. This suggests that although non-strategic players' choices are not strictly i.i.d., they are highly noisy.

4.2.2 Out of Sample Predictions: From Beauty Contest to the Money Request Game

To test whether our PRAI classification from the beauty-contest games predicts behavior in the 11-20 game, we examine each group separately. Figure 2 presents the empirical cumulative distribution functions (CDFs) for each of our six PRAI levels alongside the equilibrium CDF. Although players at all PRAI levels deviate from the equilibrium strategy, those with higher index values come closer to it. To formally test this relationship, we estimate a probit regression where the dependent variable equals 1 if the subject chooses an action not part of the mixed-strategy Nash equilibrium. The results in Table 3 confirm

Table 2: Correlations of Choices: Strategic vs Non-Strategic Players

Strategic Players					
	Choice 1	Choice 2	Choice 3	Choice 4	Choice 5
Choice 1	1.000				
Choice 2	0.696	1.000			
Choice 3	0.705	0.790	1.000		
Choice 4	0.648	0.721	0.745	1.000	
Choice 5	0.666	0.740	0.744	0.782	1.000
Non-Strategic Players					
	Choice 1	Choice 2	Choice 3	Choice 4	Choice 5
Choice 1	1.000				
Choice 2	0.358	1.000			
Choice 3	0.214	0.280	1.000		
Choice 4	0.217	0.153	0.283	1.000	
Choice 5	0.277	0.199	0.241	0.371	1.000

Interpretation: The correlation coefficient between the values chosen the first and second time the subjects played games with same value of m is 0.696.

that the probability of selecting such an action decreases with our index.

Table 3: Probit Regression of Out-of-Equilibrium Actions

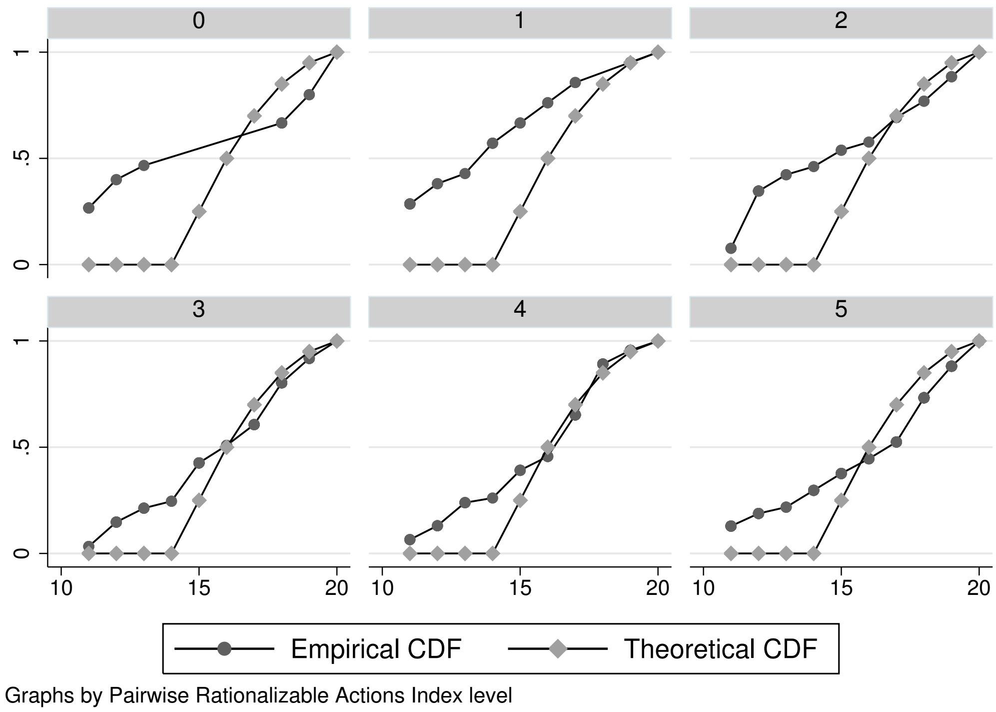
Variable	**Coefficient**	**Standard Error**
PRAI level	-0.125*	0.052
Intercept	-0.021	0.195
Number	270	
Log-likelihood	-167.473	
$\chi^2_{(1)}$	5.671	

Notes: Significance levels : * = 5 %. The Pairwise Rationalizable Actions Index (PRAI) quantifies the consistency of players' choices across five pairs of games. Players are distributed into six levels: an index value of 0 indicates that subjects systematically choose a higher number when $m = \frac{2}{3}$ than when $m = \frac{4}{3}$, while an index of 5 indicates no such inconsistency.

Moreover, the cumulative distribution of strategies chosen by low-level players does not differ significantly from a uniform distribution, whereas that of high-level players (i.e., PRAI levels 4 or 5) does. The Chi-squared test against the discrete uniform distribution over $\{11, 12, \ldots, 20\}$ yields a p-value of 0.22 for low-level players and 0.01 for high-level players.

While strategic sophistication tends to be unstable across different categories of games

Figure 2: Cumulative Density Functions (CDF) in the 11-20 Game



(see, e.g., Georganas et al. (2015) and Engelmann & Hollard (2010)), our analysis shows that the propensity for random choice persists across games. Such out-of-sample predictions can only be performed in environments that elicit a substantial fraction of random players. For example, matrix games are less likely than abstract games like the beauty contest to generate many identifiable reasonless choices.

5 Discussion on the Origins of Reasonless Choices

Collected evidence allows us to validate some assumptions, and discard others, regarding the origins of reasonless choices. We here examine each in turn.

5.1 Are Non-Strategic Players Simply Not Paying Attention?

It is challenging to determine whether lab subjects are paying attention or exerting effort, but our data provide some insight. We recorded reaction times in the beauty-contest games as an indicator of cognitive effort. Subjects who avoid effort or attention tend to respond faster than those who spend time developing strategic choices. Conversely, if subjects are distracted by other factors, their reaction times will also differ from those of the most strategic players.

First, non-strategic players spend slightly *more* time to make their choices than other players. On average, non-strategic players spend 28.87 seconds per decision compared to 25.38 seconds for strategic players. The difference is statistically significant (p-value = 0.059).[11] This difference also holds for the 11-20 game in Phase 2, as level-0 players also spend more time than others (196.8 seconds vs 175.1; p-value=0.029).

The second, and most surprising, results concerns reaction times as the parameter m changes. Typically, subjects play progressively faster over successive games. However, at some point they are randomly confronted with a change in the parameter m. For instance, after three consecutive games with $m = \frac{2}{3}$, a game with $m = \frac{4}{3}$ follows. When subjects are confronted with this change, they need adjust, which increases their reaction time relative to the previous game. One might expect that non-strategic players—if inattentive—would not react to such changes since their strategy does not vary with m. Yet, our data show that non-strategic players do respond to changes in m. A fixed-effects regression of reaction time on a dummy for the first change in m and its interaction with our PRAI classification, controlling for period, shows that facing the first change increases reaction time by 5.39 seconds (p-value = 0.066) for strategic players, with no significant difference for non-strategic players (interaction coefficient = 1.16, p-value = 0.802).[12] This evidence suggests that non-strategic players are aware of the change, even if they fail to adjust their strategy accordingly.

5.2 Are the Stakes Too Low to Motivate Effort?

Our experiment lasted about 20 minutes and subjects were already on site (mostly socializing). They received on average about \$16 (11€) for these 20 minutes. Since no transport costs were incurred, this equates to an hourly wage of \$48 (33€), which is high compared

[11]The recorded time for the first decision includes the time taken to read the instructions. We thus do not have a reliable measure of reaction time for the first decision.

[12]The period in which the first occurs is similar for both player types (2.53 vs 2.62, p-value=0.43).

to the average hourly wage and slightly above the rate in typical lab experiments.

If stakes drive reasonless choices, we would expect fewer such choices when stakes are higher. Our design does allow us to test that assumption in two ways. First, our experiment includes two treatments for the beauty contest game: one with two players and another with three players. Although the winner's reward is constant, subjects in the two-player game have a higher expected payoff (50 points, i.e., 10€) compared to those in the three-player game (33.3 points, i.e., 6.66€). We calculate our index separately for these two groups and find that 45.52% of subjects are classified as non-strategic in the two-player version versus 45.59% in the three-player version. Thus, a 50% rise in expected payoff does not affect the non-strategic percentage.

Second, in the final phase of the experiment, subjects have the chance to win $200, after having received a cash payment (and feedback). Even then, non-strategic players hardly modify their strategies.

In sum, the frequency of reasonless choices does not change with stakes, suggesting that stakes are not a key driver of such behavior.

5.3 Are Non-Strategic Players Unable to Think Strategically?

The purpose of recruiting chess players during an international tournament was to eliminate the possibility that subjects simply lack strategic thinking. Chess requires players to anticipate their opponents' moves. We are therefore confident that all subjects, including those classified as non-strategic, are capable of strategic thinking. Surprisingly, Elo rating does not seem to correlate with our classification: non-strategic players have a mean Elo of 1768 (SD = 30), while strategic players have a mean Elo of 1814 (SD = 27), a difference that is not statistically significant (p-value = 0.25). Thus, being strong at chess does not necessarily translate into strategic behavior in experimental games (see Levitt et al. 2010).

5.4 Does it Pay to Use Reasonless Choices?

Playing randomly might be a deliberate strategy. In some contexts, such as in hide-and-seek, random play is optimal and can yield high gains, despite not being rational. Some

Table 4: Earnings in Euros by PRAI Level

Index Value	**Earnings** Mean	Std. Dev.	**P-Value**
0	3.6	1.9	
1	4.7	1.7	0.109
2	5.1	2.0	0.451
3	5.5	1.8	0.287
4	5.4	2.2	0.773
5	7.2	2.7	0.000

Notes: P-values reflect t-tests comparing earnings at each index level to the previous level. Std. Dev. stands for Standard Deviation.

subjects may be "street smart" adopting non-optimal yet effective strategies in a bounded rationality setting. We examine this possibility by analyzing how earnings vary with the frequency of reasonless choices in the beauty-contest games. Table 4 shows that earnings actually increase nearly monotonically with the PRAI index: players with an index of 0 earn only half as much as those with an index of 5 (3.6 vs. 7.2). In our experiments, reasonless choices do not pay. We can thus rule out the possibility that many subjects use unknown yet effective heuristics.[13] If such unusual yet effective strategies were used by more than a small fraction of players, we would have observed smaller differences in earnings across groups.

5.5 Poor Instructions?

The money request game features concise instructions—just a few lines on the screen—and subjects must answer four comprehension questions correctly before proceeding. These questions present hypothetical scenarios and ask whether the indicated payoffs are correct. Table 5 shows that non-strategic players require significantly more attempts to answer these questions correctly, making roughly two mistakes on average (median = 6 attempts), while most strategic players answer correctly on their first try (median = 4 attempts). Differences in both the mean and median are significant at the 1% level.

We examined reaction times and error rates for each comprehension question. Non-

[13] For example, recent evidence suggests that a fair proportion of subjects in the guessing game–a game similar to the beauty contest–employ rules not described in standard game theory but that nevertheless make sense (see Fragiadakis et al. (2016)).

Table 5: Attempts Required to Correctly Answer Four Comprehension Questions

Level	**Number of Attempts**		
	Mean	Median	Std. Dev.
High PRAI (strategic)	4.74	4	1.09
Low PRAI (non-strategic)	5.67	6	1.40

Notes: The Pairwise Rationalizable Actions Index (PRAI) quantifies the consistency of players' choices across five pairs of games. We define players with a PRAI below 3 as non-strategic and those with a PRAI of 4 or above as strategic. Std. Dev. stands for Standard Deviation.

strategic players spend significantly more time thinking than strategic ones (18s on average vs. 14s, $p < 0.01$), and their error rates vary by question. For example, their success rate increases significantly from 60% on the first question to 72% on the second ($p = 0.037$).

Direct observation and exit interviews did not reveal any subgroup-specific attitudes.We controlled for language issues by screening subjects during recruitment and verifying their international chess player identification (FIDE), which confirmed that 88% were French nationals while the remainder were fluent in French or English. A chi-squared test of independence between French nationality and the occurrence of low PRAI indicates no statistically significant association $p = 0.360$, and a two-sample t-test on the number of attempts shows no dependence on nationality ($p = 0.337$); restricting the analysis to French subjects produces similar results.[14]

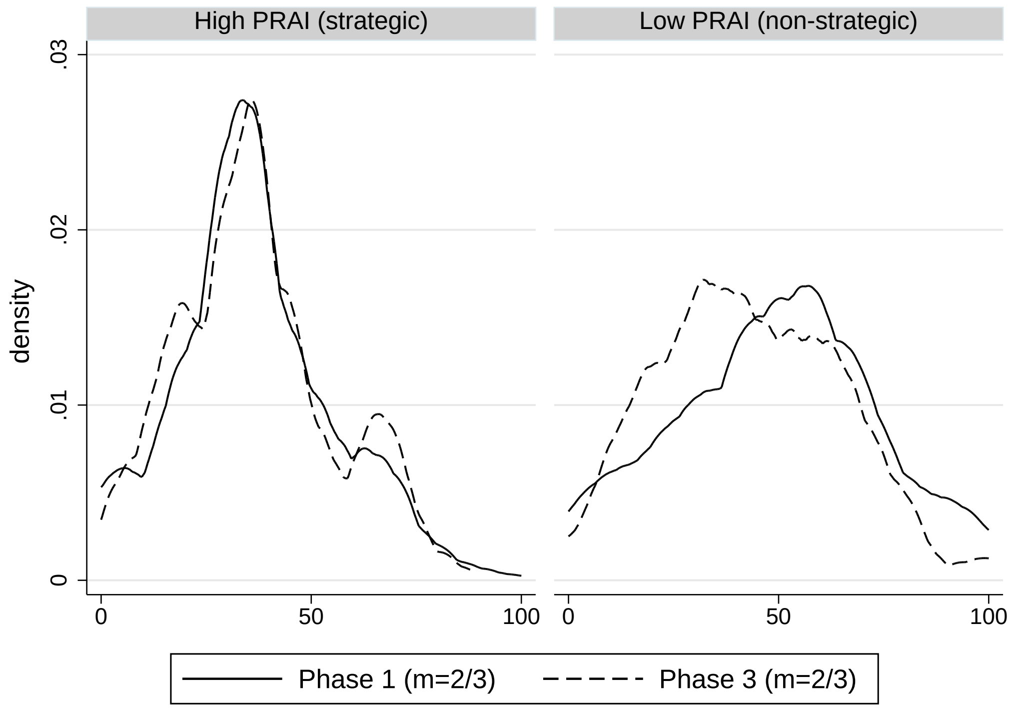
Overall, our controls indicate that non-strategic players are genuinely trying their best. Evidence suggests that while they recognize the need to act, they fail to determine the correct approach.

5.6 Slow Learners?

The previous section provides evidence that subjects who struggle with comprehension questions are more likely to make reasonless choices. We further examine Phase 3, before which players receive payment and feedback, to assess whether the opportunity to learn affects behavior. Overall, feedback appears to have a limited effect (Figure 3): when $m = \frac{2}{3}$, subjects' average play drops slightly from 42 in Phase 1 to 39.4 in Phase 3. The graphs reveal that strategic players continue to use similar strategies as in Phase 1, while

[14]Results are available upon request.

Figure 3: Actions in Phase 1 ($m = 2/3$) and Phase 3 ($m = 2/3$)



Notes: The Pairwise Rationalizable Actions Index (PRAI) quantifies the consistency of players' choices across five pairs of games. We define players with a PRAI below 3 as non-strategic and those with a PRAI of 4 or above as strategic. Std. Dev. stands for Standard Deviation.

non-strategic subjects show only modest improvement. Even when comparing games that differ in key aspects (e.g., number of players and stakes), the two groups remain significantly different (t-test, $p = 0.0139$; Kolmogorov-Smirnov test, $p = 0.004$). In sum, players making reasonless choices appear to be slow learners.They need more trials to answer the control questions correctly in Phase 2. They seem to struggle to navigate unfamiliar abstract environments.

5.7 Are Some Subjects Cognitively Challenged?

Burnham et al. (2009) find that players with low IQ are far more likely to play dominated strategies in the beauty-contest game and be classified as level-0, suggesting that this level may be a stable individual trait across games. However, Brañas-Garza et al. (2012) show that while the Cognitive Reflection Test (CRT) is associated with an identifiable pattern in the beauty-contest game, the Raven test is not. In addition, Georganas et al. (2015)

elicit five measures of cognitive ability to predict behavior (IQ quiz, the Eye Gaze quiz, a memory quiz, the CRT, and a one-player Takeover Game). However, correlations among these measures are weak. Thus, even if poor cognitive ability contributes to reasonless choices, no straightforward measure reliably predicts them.

6 Conclusion

We have shown that reasonless choices are a robust phenomenon, consistently observed across numerous experiments. As a scientific fact, their existence merits dedicated study. By reviewing existing evidence and adding new empirical observations, we support the assumption that reasonless choices are primarily made by subjects who fail to "solve" games during the experiment. Reasonless choices are not driven primarily by stakes, instructions, or overall strategic ability. Rather, when some players act randomly, they tend to randomize across available strategies with little strategic consideration.

Attitude toward reasonless choices warrant clarification. On one hand, if an experiment's primary goal is to test theories, it may be acceptable to introduce a selection bias that limits reasonless choices. After all, adding more random players only adds noise in the data, which provides little insight into the relative explanatory power of considered theories. Note that, even high-performing students sometimes make reasonless choices; although their frequency can be reduced, they cannot be entirely eliminated. On the other hand, when external validity is an important concern, selecting subjects becomes an issue. Students at top universities may not represent the broader population, which is likely to be less educated and include a higher incidence of reasonless choices.

References

Agranov, M., Caplin, A. & Tergiman, C. (2015), 'Naive play and the process of choice in guessing games', *Journal of the Economic Science Association* **1**, 146–157.

Agranov, M., Potamites, E., Schotter, A. & Tergiman, C. (2012), 'Beliefs and endogenous cognitive levels: An experimental study', *Games and Economic Behavior* **75**, 4449–4463.

Alekseev, A., Charness, G. & Gneezy, U. (2017), 'Experimental methods: When and why contextual instructions are important', *Journal of Economic Behavior & Organization* **134**, 48–59.

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

Brañas-Garza, P., García-Muñoz, T. & González, R. H. (2012), 'Cognitive effort in the beauty contest game', *Journal of Economic Behavior & Organization* **83**(2), 254–260.

Buhren, C. & Frank, B. (2012), 'Chess players' performance beyond 64 squares: A case study on the limitations of cognitive abilities transfer', *Talent Development and Excellence* **4**(2), 157–169.

Buhren, C., Frank, B. & Nagel, R. (2012), 'A historical note on the beauty contest'. MAGKS Papers on Economics 201211.

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

Burnham, T. C., Cesarini, D., Johannesson, M., Lichtenstein, P. & Wallace, B. (2009), 'Higher cognitive ability is associated with lower entries in a p-beauty contest', *Journal of Economic Behavior and Organization* **72**(1), 171–175.

Camerer, C., Ho, T.-H. & Chong, J.-K. (2004), 'A cognitive hierarchy model of games', *Quarterly Journal of Economics* **119**(3), 861–898.

Chen, C.-T., Huang, C.-Y. & Wang, J. T.-y. (2018), 'A window of cognition: Eyetracking the reasoning process in spatial beauty contest games', *Games and Economic Behavior* **111**, 143–158.

Chou, E., McConnell, M., Nagel, R. & Plott, C. R. (2009), 'The control of game form recognition in experiments: Understanding dominant strategy failures in a simple two-person "guessing" game', *Experimental Economics* **12**, 159–179.

Coricelli, G. & Nagel, R. (2009), 'Neural correlates of depth of strategic reasoning in medial prefrontal cortex', *Proceedings of the National Academy of Sciences* **106**(23), 9163–9168.

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

Dohmen, T. & Jagelka, T. (2024), 'Accounting for individual-specific reliability of self-assessed measures of economic preferences and personality traits', *Journal of Political Economy Microeconomics* **2**(3), 399–462.

Engelmann, D. & Hollard, G. (2010), 'Reconsidering the effect of market experience on the "endowment effect"', *Econometrica* **78**(6), 2005–2019.

Enke, B. & Graeber, T. (2023), 'Cognitive uncertainty', *Quarterly Journal of Economics* **138**(4), 2021–2067.

Enke, B., Graeber, T., Oprea, R. & Yang, J. (2024), Behavioral attenuation, Technical report, National Bureau of Economic Research.

Fragiadakis, D. E., Knoepfle, D. T. & Niederle, M. (2016), ‘Who is strategic?’.

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

Grosskopf, B. & Nagel, R. (2008), ‘The two-person beauty contest’, *Games and Economic Behavior* **62**, 93–99.

Ivanov, A., Levin, D. & Niederle, M. (2010), ‘Can relaxation of beliefs rationalize the winner’s curse? an experimental study’, *Econometrica* **78**(4), 1435–1452.

Levitt, S., List, J. & Reiley, D. H. (2010), ‘What happens in the field stays in the field: Professionals do not play minimax in laboratory experiments’, *Econometrica* **78**(4), 1413–1434.

Levitt, S., List, J. & Sadoff, S. (2009), ‘Checkmate: Exploring backward induction among chess players’, *American Economic Review* **101**(2), 975–990.

Lopez, R. (2001), ‘On p-beauty contest integer games’. UPF Economics and Business Working Paper 608.

Palacios-Huerta, I. & Volij, O. (2009), ‘Field centipedes’, *American Economic Review* **99**(4), 1619–1635.

Weber, R. A. (2003), “learning’with no feedback in a competitive guessing game’, *Games and Economic Behavior* **44**(1), 134–144.

Yechiam, E. & Zeif, D. (2023), ‘Revisiting the effect of incentivization on cognitive reflection: A meta-analysis’, *Journal of Behavioral Decision Making* **36**(1), e2286.

Appendices

A Simulating our Pairwise Rationalizable Actions Index (PRAI) for a Homogeneous Population

We assess whether the observed PRAI distribution could have arisen by chance from a homogeneous population of random players. We assume that every player chooses actions randomly from the joint empirical distribution of actions. For each simulation run, we generate 270 individuals. For each individual, we draw five pair of actions. Each pair is drawn from the empirical joint distribution of pairs of actions against each type of opponent (i.e., A, B, C, D or Random). We then calculate the proportion of simulated players at each level of our index. Over 9999 simulation runs, we report the mean and the 1st and 99th percentiles of these proportions in Table 6.

Table 6: Simulated versus actual proportions

	Simulated proportions			
Index Value	1st percentile	Mean	99th percentile	Observed proportion
0	0	0.2	1.1	5.6
1	0.7	2.8	5.6	7.8
2	8.5	13.2	18.1	9.6
3	24.4	30.9	37.8	22.6
4	29.3	36.0	42.6	17.0
5	11.9	16.8	22.2	37.4
Total	-	100.0	-	100.0

As shown in Table 6, the simulated proportions differ substantially from the observed data. Even under the most favorable scenarios, the proportion of level-5 players exceeds 22.2% in fewer than 1% of the 9,999 simulated draws, far below the observed figure of 37.4%. This strongly suggests that our most strategic players are not merely random players who happened to choose consistent strategies by chance. Importantly, our simulations are based on the *joint empirical distribution* of action pairs, i.e., the approach most likely to generate a simulated distribution resembling the observed one.

B Descriptive Statistics in Beauty-Contest Games

Table 7: Means and Standard Deviations for Non-Strategic Players

		$m = 2/3$		$m = 4/3$	
Opponent Type	Obs	Mean	Std. Dev.	Mean	Std. Dev.
A	123	51.6	24.3	50.1	26.6
B	123	51.6	25.1	50.7	25.5
C	123	48.6	24.7	54.7	24.3
D	123	48.3	24.8	47.4	25.8
Random	123	44.9	23.2	53.4	25.3
Overall	615	48.98	24.47	51.25	25.54

Notes: The Pairwise Rationalizable Actions Index (PRAI) quantifies the consistency of players' choices across five pairs of games. We define players with a PRAI below 3 as non-strategic. Std. Dev. stands for Standard Deviation. Opponent Type: all treatments are identical except that half of the subjects play against one opponent only, while the other half play against *two* opponents of the same level (A, B, C, D or Random). Letters denote Elo ranking (A = Elo $>$ 2150, B = 2150 $>$ Elo $>$ 1800, C = 1800 $>$ Elo $>$ 1500, D = Elo $\leq$ 1500). The higher the Elo rating, the stronger the player. "Random' indicates a device that selects strategies uniformly across the strategy space.

Table 8: Means and Standard Deviations for Strategic players

		$m = 2/3$		$m = 4/3$	
Player Type	Obs	Mean	Std. Dev.	Mean	Std. Dev.
A	147	34.8	20.5	71.0	23.7
B	147	35.0	16.9	72.3	21.3
C	147	35.9	18.6	71.4	21.3
D	147	38.1	20.9	68.7	21.7
Random	147	36.8	18.1	67.0	21.5
Overall	735	36.12	19.04	70.09	21.95

Notes: The Pairwise Rationalizable Actions Index (PRAI) quantifies the consistency of players' choices across five pairs of games. We define players with a PRAI of 4 or above as strategic. Std. Dev. stands for Standard Deviation. Opponent Type: all treatments are identical except that half of the subjects play against one opponent only, while the other half play against *two* opponents of the same level (A, B, C, D or Random). Letters denote Elo ranking (A = Elo $>$ 2150, B = 2150 $>$ Elo $>$ 1800, C = 1800 $>$ Elo $>$ 1500, D = Elo $\leq$ 1500). The higher the Elo rating, the stronger the player. "Random' indicates a device that selects strategies uniformly across the strategy space.