UNIVERSITY OF WARWICK

Economics

Rationality and Cooperation

Ali Moghaddasi Kelishomi, Daniel Sgroi and Andis Sofianos

August 2026
No: 1630

ISSN 2059-4283 (online)
ISSN 0083-7350 (print)

Warwick Economics Research Papers

Rationality and Cooperation*

Ali Moghaddasi Kelishomi[1], Daniel Sgroi[2] and Andis Sofianos[3]

[1]Loughborough University and IZA

[2]University of Warwick and IZA

[3]Durham University

August 2026

Abstract

How does rationality shape cooperation in strategic settings? We study this question in a laboratory experiment that links individual rationality, measured by consistency with the generalized axiom of revealed preference, to behaviour in an indefinitely repeated Prisoner's Dilemma. Participants are grouped by pre-measured rationality before interacting repeatedly. We find that higher rationality substantially increases cooperation and payoffs. This effect operates through a novel mechanism: more rational individuals make fewer implementation errors when executing their intended strategies, thereby sustaining cooperative outcomes. By contrast, higher cognitive ability also promotes cooperation and higher payoffs, but through a distinct channel—reducing strategic errors in responding optimally to others' actions. Our results provide the first experimental evidence linking rationality to cooperation via decision-making errors, and clarify the distinct roles of rationality and intelligence in shaping strategic behaviour. Together, the findings offer a unified account of how cognitive constraints affect cooperation in repeated games.

Keywords: Repeated Prisoners Dilemma, Cooperation, Rationality, Intelligence, Learning, Strategy Errors.

JEL classification: C73, C91, C92, D83.

*We are grateful for helpful discussions with David Gill and Eugenio Proto. We also thank seminar and conference participants at the Workshop on Decision-Making and Bounded Rationality, Loughborough University, University of Warwick. Ethical approval was provided by University of Warwick Humanities and Social Sciences Research Ethics Committee (ref HSSREC 178.24-25). The project was pre-registered with the AEA RCT Registry (https://www.socialscienceregistry.org/trials/16845). All authors contributed equally to the production of this paper and are listed alphabetically. The authors declare no conflicts of interest.

1 Introduction

Intelligence has been shown to be very important for making decisions in strategic and single-player contexts (Alaoui and Penta, 2016; Alaoui et al., 2020; Gill and Prowse, 2016; Proto et al., 2019, 2022; Sofianos, 2025; Eber et al., 2024). In particular, Proto et al. (2019) show that when subjects are matched by intelligence, only high intelligence groups converge to cooperation in an indefinitely repeated prisoner's dilemma. Proto et al. (2022) further investigate this and explain the differences between intelligence groups through implementation errors of a given strategy. Overall, the literature has shown that more intelligent players are better at identifying the more profitable course of action in a given game and additionally better at consistently following this course of action. In a non-behavioral (fully rational) context we would expect all players to be able to do this flawlessly, and so these conclusions suggest that intelligence may proxy for rationality.

In the current paper we use a more direct measure of rationality, and discover that rationality and intelligence influence behaviour in quite distinct ways. Our elicitation task follows the revealed-preference framework developed by Choi et al. (2007, 2014), but we measure rationality using the Critical Cost Efficiency Index associated with fGARP, following Polisson et al. (2020). fGARP strengthens the Generalized Axiom of Revealed Preference (GARP) for choice under risk by additionally requiring choices to respect first-order stochastic dominance. It therefore measures the extent to which a subject's choices can be rationalised by a continuous, stochastically monotone utility function. Because the resulting index captures consistency with utility maximisation across a sequence of choices, it provides a natural measure of rationality for studying behaviour in a repeated strategic setting.

In our experiment we compare how people cooperate when they are grouped according to their degree of rationality. Unlike intelligence, rationality does not follow a symmetric or normal distribution, but instead looks more like a power law, with significant clustering at lower levels. To account for this we split our sample into rationality quartiles, and then match them within each quartile where they play the indefinitely repeated Prisoners Dilemma game. We then directly

compare how the highest quartile (the group of high rationality types) compare to the lowest three quartiles (lower levels of rationality types).[1] As well as measuring rationality, we also measure intelligence via a standard IQ test (John and Raven, 2003), hierarchical reasoning via the 11-20 game (Arad and Rubinstein, 2012) which has been shown to trigger level-k reasoning (Costa-Gomes et al., 2001), alongside a slew of demographic control variables.

Before explaining our results we need to first provide some explanation of the different types of error that we examine in our paper, which mirror the error types used in Proto et al. (2022), allowing us to make direct comparisons. We also need to discuss the relationship between intelligence and rationality.

Intelligence and rationality are certainly correlated, but that correlation is far from perfect. Quoting from Eber et al. (2024): "...we expect the behavior of people with higher cognitive abilities to be closer to what is predicted by standard rational models, compared to the behavior of people with lower cognitive abilities". However, the two concepts are not perfect substitutes. Much has been written on the distinction between intelligence and rationality in what has become known within psychology and philosophy as "the Great Rationality Debate", neatly summarized in Stanovich (2012) from which we draw another quote: "There are individual differences in rational thinking that are less than perfectly correlated with individual differences in intelligence because intelligence and rationality occupy different conceptual locations in models of cognition". Our own work takes these two points as a (testable) starting point. Our findings suggest that intelligence and rationality are indeed correlated but only mildly, thus, neatly supporting both propositions. We next show that rationality matters a great deal when making decisions in the indefinitely repeated Prisoners Dilemma, which mirrors earlier work on the importance of intelligence in the same context (Proto et al., 2019, 2022). Our next finding draws on exactly the point made by Stanovich (2012) and sharpens his notion of "different conceptual locations" in a sense perhaps more meaningful for play in games. We find that being more rational means making fewer errors when implementing a chosen strategy (what Proto et al.,

[1] Since we had prior knowledge of the distribution of rationality from earlier studies on rationality such as Choi et al. (2014); Moghaddasi Kelishomi and Sgroi (2024) and our pilot experiment, we pre-registered this exact level of comparison. See Appendix B where we report the distribution of this measure from some earlier studies including the pilot data for the present experiment.

2022, call action errors, ϵ_A) and note that this mechanism is very different from the mechanism through which intelligence operates in Proto et al. (2022), which works instead mainly through errors relating to optimal response for a chosen strategy (what they call transition errors, ϵ_T). In fact, if rationality operates through action errors and intelligence through transition errors, but intelligence is mildly correlated with rationality, this also explains why there is also a mild relationship between intelligence and action errors when rationality is not measured separately. Taking this altogether we can not only better explain the role of rationality, but also support the idea that rationality and intelligence should not be viewed as substitute concepts, rather they measure very different aspects of decision-making in an integrated error-model of behavior.

Our finding that high-rationality subjects make significantly fewer action errors (ϵ_A) is rooted in the interpretation of rationality as a measure of procedural discipline. As defined in our strategy estimation framework, action errors represent implementation failures or execution-level lapses where a player fails to choose the action prescribed by their current strategic state. Adherence to the fGARP axiom, as elicited in the budgetary choice task, provides a direct proxy for the cognitive capacity required to maintain such consistency. As noted by Choi et al (2014), violations of GARP typically result from "trembles" or the strategic "simplification" of complex decision rules, reflecting a lack of the implementation focus required for high-quality decision-making.

This link between rationality and implementation consistency finds robust support in the experimental evidence of Nielsen and Rehbeck (2022). Their design reveals a pervasive execution gap: they find that a subject's ex-ante preference for a rational rule does not predict their actual adherence to it, with subjects frequently violating axioms they explicitly wish to follow. This distinction between strategic intent and execution capacity validates our interpretation that implementation is a separate cognitive challenge from strategic reasoning. Furthermore, their finding that participants reconcile 47% of inconsistencies in favor of the rational rule, while only 13% renounce the rule, provides empirical confirmation that these deviations are implementation failures rather than stable behavioural preferences. Consequently, our results suggest that the fGARP-based CCEI identifies the specific individual trait that enables high-rationality

types to close this execution gap and sustain cooperation by minimizing implementation-level noise.

Our emphasis on action errors (ϵ_A) as a primary driver of strategic failure aligns with the recent findings of Guan and Oprea (2026), who identify "implementation complexity" as a first-order determinant of strategy selection. They define implementation complexity as the subjective cognitive effort required to execute a chosen rule round-by-round and find that subjects exhibit a profound aversion to it, frequently sacrificing up to 44% of their potential earnings simply to avoid complex implementation tasks. This finding provides a powerful empirical justification for our focus on the "execution gap" in repeated games; it suggests that the round-by-round act of following a rule is not a trivial operation but a costly cognitive hurdle that participants will pay significantly to avoid. By identifying individual rationality (measured by fGARP) as the specific trait that enables subjects to minimize these implementation inconsistencies, we ground our results in a framework where the CCEI serves as a direct measure of the capacity to overcome these subjective implementation costs. Furthermore, Guan and Oprea (2026) observe that this complexity aversion is robust to variation in cognitive ability, which provides further external support for our unified account of strategic constraints: while intelligence governs the "hardware" of tracking states and transitions (ϵ_T), rationality governs the procedural discipline required for the consistent execution of actions (ϵ_A).

The rest of the paper is organized as follows: Section 2 presents the experimental design. Section 3 summarizes the key results on behavior within the experiment. Section 4 summarizes a framework allowing us to make a functional difference between errors that are attributable to variations in intelligence, and errors that are attributable to variations in rationality. We follow in Section 5 by estimating the error parameters, and comparing those findings to earlier work on intelligence and errors in strategic decision-making. Section 6 concludes. The Supplementary Appendix contains further details on the experimental design, descriptive statistics and additional tables and figures.

2 Experimental Design

Our experimental design mirrors the framework used in Proto et al. (2019, 2022). The key difference is that instead of grouping participants according to intelligence, we group participants according to their degree of rationality when asked to play the indefinitely repeated Prisoner's Dilemma game. Additionally, in order to mitigate any possible attrition issues and simplify the logistics of implementation we administer the experiment within a single session instead of the two-day paradigm implemented in Proto et al. (2019, 2022).

2.1 Measuring participant characteristics

In the first part of the session, participants complete a series of tasks that elicit different types of abilities. These were sequenced in all sessions in the order we describe them below.

Intelligence. Participants first complete a Raven Advanced Progressive Matrices (APM) test consisting of a sequence of 12 questions. The participants are allowed a maximum of 10 minutes for the entire test during which they can move forwards and backwards and change their answers within this time limit. We use Set II of the APM and the questions are presented to the participants in a sequence of increasing difficulty. For each question, a 3×3 matrix of images is shown on the screen, with the bottom-right image missing. The participants are then asked to complete the pattern by selecting one out of 8 displayed images shown on their screen. This test is a non-verbal test that measures fluid intelligence and has been widely used in the economics literature (Some examples: Basteck and Mantovani, 2018; Charness et al., 2018; Gill and Prowse, 2016; Proto et al., 2019, 2022). Participants earn £1 per correct answer out of 2 randomly selected answers out of the 12.

Rationality. The second task the participants complete is a measure of decision-making quality based on the experimental framework developed by Choi et al. (2007, 2014). This task utilises an intuitive graphical interface to elicit a rich dataset of individual choices from a series of independent decision problems under risk. The central aim of this budgetary choice proce-

dure is to test the consistency of observed behaviour with the utility maximization hypothesis. Participants are presented with a series of 25 independent decision problems under risk. Each problem requires a choice from a two-dimensional budget line, representing an allocation of points between two accounts, A and B, which correspond to the horizontal and vertical axes, respectively.[2] Each account is equally likely to be selected for payment, meaning participants receive the points allocated to one of the accounts with a 50% probability. To record their choices, participants first click a point on the budget line and can then use a slider to precisely move their selection to their desired allocation.

Participants are provided with extensive instructions and 2 practice rounds to familiarize themselves with the graphical interface before proceeding to the 25 payment-relevant rounds. Each point earned in this task is converted at a rate where one point corresponds to £1/32.

Following the non-parametric revealed-preference approach refined by Polisson et al. (2020), we assess compliance with the *Generalized Axiom of Revealed Preference* (GARP), which is the standard behavioural requirement for utility maximization (Choi et al., 2014; Polisson et al., 2020). Specifically, we employ a strengthened version of this axiom known as fGARP (or F-GARP), which incorporates the requirement that choices respect first-order stochastic dominance (FOSD) in addition to standard revealed-preference cycles (Polisson et al., 2020). This test determines whether a subject's behaviour can be rationalized by a continuous and stochastically monotone utility function.

To quantify departures from this criterion, we calculate a consistency score based on the Afriat (1972) Critical Cost Efficiency Index (CCEI). We utilise the CCEI associated with fGARP compliance following the procedure suggested by Polisson et al. (2020). This index measures the extent by which all budget constraints would need to be shifted (or contracted) to remove every violation of the fGARP axiom for a given participant. The CCEI has a range between 0 and 1, where a value of 1 indicates perfect compliance with stochastically monotone utility maximisation. Values closer to 1 indicate that only a small shift is necessary to obtain choices that

[2]In each round, the computer selects a budget line that intersects both axes at points between 40 and 400, ensuring that at least one of the intercepts is at or above 200 points.

would satisfy the axiom, whereas lower values quantify more severe violations. Put differently, this CCEI provides a non-parametric measure of a subject's overall consistency with stochastically monotone preferences, often interpreted as the fraction of "wasted income" attributable to decision-making errors. Following Choi et al. (2014), this task provides a measure of *rationality*, defined by the degree of consistency between observed choices and the utility maximisation hypothesis.

Our choice of the Afriat CCEI based on the FGARP (or fGARP) axiom, following the methodology of Polisson et al. (2020), is motivated by both theoretical stringency and empirical robustness. While the standard GARP requires only a complete and transitive preference ordering, it can be insensitive to choice patterns that violate First-Order Stochastic Dominance (FOSD). For example, a subject who consistently chooses a corner solution, allocating all points to a single account regardless of price, would satisfy standard GARP with a CCEI of 1.0, despite failing to choose the cheaper state-contingent asset (Choi et al., 2014). By utilising FGARP, which applies the revealed preference test to a combined dataset of actual and mirror-image observations, we impose the normatively appealing requirement of stochastically monotone utility maximistion. This ensures that strategical simplification or inattentiveness to relative prices is correctly identified as a reduction in decision-making quality.

We further address potential concerns regarding the use of the Afriat-based index over alternative measures such as Varian's inconsistency index. While Varian's index is theoretically more discriminating, allowing for observation-specific budget adjustments, it is computationally demanding (NP-hard) and often requires approximations for larger datasets (Polisson et al., 2020). In contrast, the Afriat Critical Cost Efficiency Index (CCEI) has seen widespread adoption in the empirical literature for quantifying departures from rationality. This prominence is driven by its intuitive economic interpretation as a measure of wasted expenditure and its high degree of computational efficiency. These practical advantages have made the CCEI an increasingly popular tool for evaluating individual decision-making quality across various domains (see, for instance, Choi et al., 2007, 2014; Carvalho et al., 2016; Stango and Zinman, 2023; Moghaddasi Kelishomi and Sgroi, 2024).

To verify that our results are not sensitive to this choice of index, we conducted a counterfactual analysis. Because existing procedures for Varian's index, such as the code provided by Halevy et al. (2018), are limited to the GARP axiom, we compared quartile assignments derived from Afriat-GARP with those from Varian-GARP (using both Mean and SSQ aggregators). A detailed breakdown of these robustness checks, including the full correlation matrices and quartile transition probabilities, is provided in Appendix A. This exercise confirms that the mappings between these measures are exceptionally close: we find over 94% agreement in quartile assignments, with Kappa values exceeding 0.90 in all cases. Only 3-6% of participants shift quartiles, and exclusively by a single rank, with no large reassignments observed.[3] This high rank correlation is further supported by Polisson et al. (2020), who developed a new procedure to calculate Varian indices for FGARP using the data from Halevy et al. (2018). They report a correlation of 0.975 and a rank correlation of 0.998 between the Afriat and Varian versions of the FGARP measure. This high concordance confirms that the Afriat CCEI remains a robust and reliable proxy for individual rationality in our experimental setting.

Strategic Sophistication. The third task is a measure of strategic sophistication. We administer the 11-20 game (Arad and Rubinstein, 2012), where participants are randomly matched in pairs and each requests an amount of money. The amount requested must be an integer between 11 and 20. Each participant receives the amount they request, while they can receive an additional amount of 20 if they request exactly one unit less than their matched partner. This game is a quick and simple way to elicit a measure of iterative thinking or level-k reasoning of individuals. Requesting 20 seems natural and salient and has been typically seen as the level-0 choice in this game. Naturally, the best response to the level-0 choice would be a request for 19, while the same logic applies for the best response to 19 and so on. Alaoui and Penta (2016) implement a modified version of this game to test their model, while Goeree et al. (2018) introduced a graphical version of the game. More recently, Alós-Ferrer and Buckenmaier (2021)

[3]The reported correlations and concordance statistics are calculated based on the subsample of participants (N=196) for whom an exact solution for Varian's index could be obtained. As noted by Polisson et al. (2020), the computational procedure developed by Halevy et al. (2018) is NP-hard and may not deliver exact values for all subjects in a given sample, sometimes resulting in approximations. To ensure maximum precision in our comparison between the Afriat and Varian frameworks, we exclude these approximated cases from the analysis.

Table 1: **PD Stage Game.** *C*: Cooperate, *D*: Defect.

	C	D
C	(48, 48)	(12, 50)
D	(50, 12)	(25, 25)

also implement this game to understand sophistication using a number of variants. Each point earned in this task corresponds to £0.10.

2.2 The Prisoner's Dilemma Game

In the second part of the session, participants play an induced infinitely repeated Prisoner's Dilemma (PD) game. Participants are split in four groups according to their rationality level. In previous studies of intelligence and strategic interactions, the split was done at the median. Since rationality elicited using the Choi et al. (2014) task does not follow a symmetric or normal distribution and entails substantial clustering at lower levels of scores, we split participants into rationality quartiles instead. Specifically, participants are ranked according to their CCEI score and then divided into four quartiles. Participants play the indefinitely repeated PD game within their quartile group. Note that participants are not explicitly informed of matching taking place within quartile groups. Specifically, participants are informed that they will be matched with someone in the room (i.e., lab).

The stage game is presented in Table 1. The payoffs are reported in terms of experimental currency units (ECU), each unit corresponds to £0.004, and participants receive the sum off all earnings across all rounds that are played.

To induce infinite repetition, participants get matched and continue playing with the same partner according to a continuation probability, specifically $\delta = 0.75$. We use a pre-drawn realisation of the continuation probability to ensure comparability across all sessions.[4] We refer to each repeated game played as a supergame and to the round within a specific supergame as a period. Finally, we refer to round as the overall count of times the stage game is played across

[4]The pre-drawn realisation is the same as the one used in Proto et al. (2022) and Lambrecht et al. (2024). In Table A8 in the appendix, we list the length of each supergame.

supergames in a given session. Play continues until the completion of the 30th supergame (which entails 98 rounds). This termination rule is not disclosed to the subjects as is commonly done in similar implementations of indefinite repetition in the lab.

The matching of partners happens within a quartile group under an anonymous and random re-matching protocol. Partners remain matched while the random continuation rule determines the match to be continued. When the match is terminated, participants are re-matched with someone else within their quartile group. We implement an approximately perfect-strangers matching protocol within quartile groups. That is, we ensure that all match combinations are exhausted before participants are re-matched with anyone they have faced previously. Each round is completed once every participant in a session has submitted their choice. At the end of each decision round, the screen summarises the outcome of the given decision round by indicating the participants' own decision, their partner's decision, and the ECU they earned for that particular round.

2.3 Other Measures

Once the Prisoner's Dilemma game was completed, we elicited a few other measures from the participants. We elicit risk attitudes using an incentivised multiple price list task (Holt and Laury, 2002), a hypothetical multiple price list for time preferences, Big-Five personality traits using the 44 questions (with answers coded on a Likert scale) developed by John et al. (1991), and some socio-demographic questions. The full demographics questionnaire is included in the instructions in the appendix.

2.4 Implementation

All sessions were held in October 2025 in the CeDEx lab at the University of Nottingham. A total of 256 subjects took part in 8 sessions. They earned on average £23 and sessions lasted 90 minutes. The experiment was programmed using Otree (Chen et al., 2016).

All experimental instructions and further details are included in appendix F.

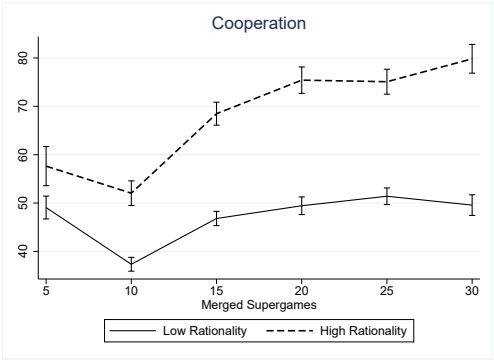
3 Experimental Results

We begin the analysis by comparing cooperation rates and payoffs, and their evolution over time, between the high-rationality and low-rationality groups. As noted above, throughout the analysis we pool rationality quartiles 1-3 into a single low-rationality group and define quartile 4 as the high-rationality group. Importantly, this pooling is performed at the level of analysis only.[5] In the experiment, subjects were matched exclusively with others from the same rationality quartile, so that interactions always took place within homogeneous quartile-specific sessions (i.e. quartile 1 subjects interacted only with quartile 1 subjects, and similarly for quartiles 2 and 3). For the purposes of the main analysis, however, we aggregate outcomes across quartiles 1-3. More fine-grained comparisons across all four rationality quartiles are reported in Appendix C.

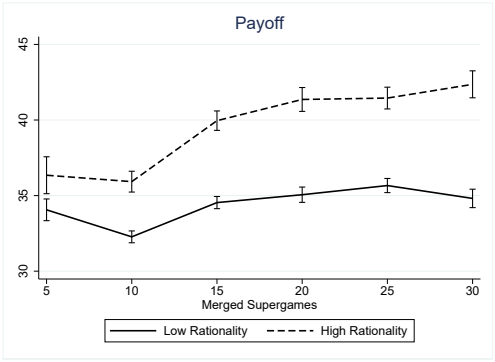
In Figure 1 we present average cooperation (panel a) and average payoff (panel b) between the low and high rationality groups. Clear differences in both cooperation and payoffs are evident between the two groups. An initial gap is already present in the first five supergames, with high-rationality subjects exhibiting higher cooperation rates and earning higher payoffs. As play progresses, this gap widens substantially. By the later supergames, the average cooperation rate of the high-rationality group approaches 80 per cent, compared to approximately 50 per cent for the low-rationality group. A closely analogous pattern is observed for average payoffs, as shown in the right panel of the figure.

Figure 2 focuses on behaviour in the first period of each supergame, when participants are matched with new partners. In contrast to the previous figure, Figure 2 shows no difference in cooperation rates in the first five supergames between rationality groups. Over time, however, behaviour diverges markedly. High-rationality subjects increasingly choose to cooperate in the first period of new supergames, whereas cooperation among low-rationality subjects stagnates at roughly its initial level following an early decline.

[5] We do this exactly because we had prior knowledge of the substantially skewed to the left distribution of this measure that has been found in several previous implementations. See Appendix B where we report the distribution of this measure from such earlier studies.



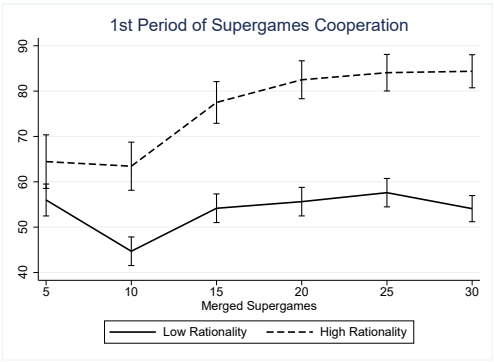
(a) Cooperation



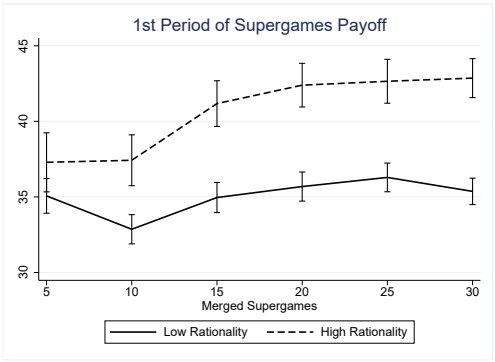
(b) Payoff

Figure 1: Average cooperation (left) and Payoff (right) computed over observations in successive blocks of five supergames of all high rationality and low rationality sessions, aggregated separately. Bands represent 95 percent confidence intervals.

This dynamic differs from the behaviour documented for intelligence in Proto et al. (2022). In that setting, cooperation among low-IQ subjects moderately increases over time, albeit more slowly and to a lower level than among high-IQ subjects. By contrast, our results suggest that lower rationality is associated not merely with lower levels of cooperation, but with a failure to sustain learning or adaptation across repeated interactions with new partners.



(a) Cooperation



(b) Payoff

Figure 2: Average cooperation (left) and Payoff (right) in the first rounds (periods) of supergames, computed over observations in successive blocks of five supergames of all high rationality and low rationality sessions, aggregated separately. Bands represent 95 percent confidence intervals.

To quantify the patterns documented in Figures 1 and 2, Table 2 reports regression estimates comparing cooperation and payoffs across rationality groups. In the first half of the supergames, subjects in high-rationality sessions earn approximately 4.2 units more and cooperate about 17% more frequently than subjects in low-rationality sessions. Columns (3) and (4) indicate that these differences increase over time: across all supergames, high-rationality subjects cooperate

around 20% more and earn roughly 5.2 additional units relative to the low-rationality group.

Table 2: Cooperation and Payoffs by Rationality Group

	Supergames < 15		All Supergames	
	(1) choice	(2) payoff	(3) choice	(4) payoff
High Rationality	0.168**	4.226**	0.208***	5.155***
	(0.058)	(1.421)	(0.049)	(1.264)
R-squared	0.061	0.130	0.089	0.183
Observations	256	256	256	256

Notes. The dependent variables are average cooperation and average payoff across all interactions. Estimates are obtained using OLS, with the low-rationality group as the omitted category. Standard errors are clustered at the session level. Results are robust to clustering standard errors at the matching-group level, where a matching group is defined by the interaction of session and fGARP quartile (32 groups in total: 8 sessions × 4 quartiles). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3 confirms that rationality does not predict cooperative behaviour in the first round of the session: neither the high-rationality dummy nor the continuous rationality measure is statistically significant. This indicates that the observed differences in cooperation across rationality groups emerge over the course of play, rather than reflecting initial behavioural differences, and are therefore consistent with learning during the session. Columns (3) and (4) show that this conclusion is robust to the inclusion of a rich set of individual controls, none of which systematically predict first-round cooperation.

Table 4 provides regression-based confirmation of the patterns illustrated in Figure 2. In the first half of the sessions, subjects in high-rationality groups become increasingly more likely to open with cooperation than subjects in low-rationality groups, indicating that differences in first-round cooperation emerge earlier among high-rationality subjects. The table also corroborates a key finding of Proto et al. (2022). As shown in columns (2) and (4), partners' past cooperative openings significantly increase the likelihood of cooperation in the first round of subsequent supergames. That is, subjects are more likely to open with cooperation when they have previously encountered partners who cooperated in the initial period, consistent with learning from past interactions.

Table 3: First Round Cooperation by Rationality and Individual Characteristics

	(1)	(2)	(3)	(4)
	Round 1 Cooperation			
FGARP Index	2.289		2.320	1.791
	(1.641)		(1.910)	(1.806)
High Rationality		1.299		1.345
		(0.254)		(0.375)
Raven (IQ)			0.907	0.899
			(0.069)	(0.067)
Strategic Reasoning			0.957	0.959
			(0.036)	(0.038)
Risk Aversion			0.950	0.945
			(0.112)	(0.112)
Patience (long-run)			0.928	0.922
			(0.082)	(0.085)
Patience (short-run)			1.129	1.127
			(0.117)	(0.116)
Female			0.908	0.925
			(0.213)	(0.217)
Openness			1.084	1.068
			(0.253)	(0.262)
Conscientiousness			1.013	1.001
			(0.213)	(0.209)
Extraversion			1.040	1.070
			(0.201)	(0.220)
Agreeableness			1.057	1.069
			(0.262)	(0.267)
Neuroticism			1.114	1.114
			(0.223)	(0.222)
Age			1.030	1.029
			(0.047)	(0.047)
Observations	256	256	256	256

Notes. Odds ratios from logit regressions are reported. The dependent variable is the choice of cooperation in the first round of each session. Standard errors are clustered at the session level and reported in parentheses. Results are robust to clustering standard errors at the matching-group level, where a matching group is defined by the interaction of session and fGARP quartile (32 groups in total: 8 sessions × 4 quartiles). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Differences in the Evolution of First-Periods Cooperation across Rationality Groups

	Supergames ≤ 15		Supergames > 15	
	(1)	(2)	(3)	(4)
	First Periods Cooperation			
High Rationality	1.213***	1.187***	1.101**	1.070
	(0.044)	(0.046)	(0.051)	(0.052)
Supergame	0.996	1.011	1.043	1.032
	(0.026)	(0.028)	(0.031)	(0.032)
Average Supergame Length	1.285**	1.439***	8.376***	8.369**
	(0.144)	(0.175)	(6.769)	(7.017)
1st Period Partner Cooperation Rates until t-1		41.999***		801939.644***
		(15.757)		(2044020.388)
Partner Cooperation Rates until t-1		1.685**		1.211
		(0.440)		(3.208)
Observations	2408	2408	1455	1455

Notes. This table reports logit estimates of cooperative choice in the first period of each supergame. The dependent variable is a binary indicator for cooperation in the first round. The low-rationality group serves as the reference category. All specifications are estimated using a fixed-effects logit model at the individual level, and coefficients are reported as odds ratios. In the later part of each session, many subjects exhibit no within-individual variation in first-round behaviour; as a result, these observations are excluded from the fixed-effects estimation in columns (3) and (4). This accounts for both the reduced sample size and the very large estimated odds ratios in these specifications. Standard errors, clustered at the session level, are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4 A Framework for Rationality, Intelligence, and Errors

This section provides a simple framework that links the two individual characteristics measured in the experiment to two distinct sources of strategic error. The framework follows the best-response adjustment model of Proto et al. (2022), but extends its interpretation to the present design. In Proto et al. (2022), a player who implements a repeated-game strategy can make two kinds of mistake: an error in action, in which the player chooses the action different from the one prescribed by the current state of the strategy, and an error in transition, in which the player fails to update the state of the automaton correctly after observing the previous-period action profile. In our setting these two errors map naturally into two different dimensions of decision making. Rationality, as measured by consistency with fGARP in the budgetary choice task, is interpreted as procedural discipline and therefore predicts the probability of action errors. Intelligence, as measured by Raven's matrices, is interpreted as the capacity to track and update the state of a strategy and therefore predicts the probability of transition errors.

4.1 Strategies as automata

Let the repeated Prisoner's Dilemma be denoted by $G = (g, \delta)$, where g is the stage game in Table 1 and $\delta = 0.75$ is the continuation probability. Let

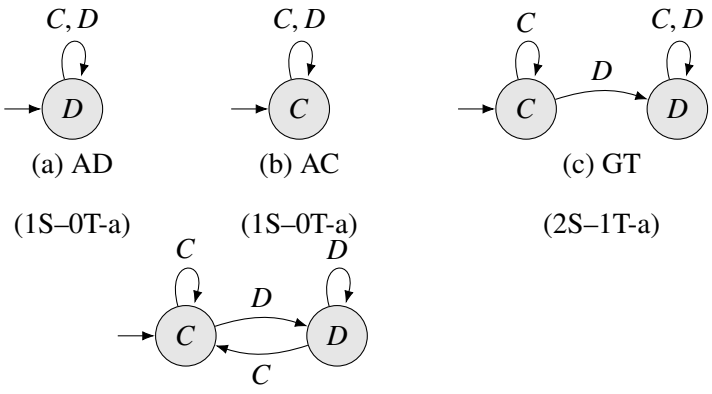
$$M = \{\text{AD}, \text{AC}, \text{GT}, \text{TFT}\}$$

be the set of strategies considered in the maximum-likelihood estimation. Each strategy $m \in M$ is represented by a finite automaton

$$m = (S_m, s_m^0, a_m, \phi_m),$$

where S_m is a finite set of states, $s_m^0 \in S_m$ is the initial state, $a_m : S_m \to \{C, D\}$ is the intended action in each state, and $\phi_m : S_m \times \{C, D\}^2 \to S_m$ is the transition rule after the realised action profile. Figure 3 gives the automata used in the estimation. The label inside each node is the action prescribed in that state. Arrows represent correct transitions. The two error probabilities

introduced below act on different parts of the diagram: action errors operate within a node by changing the realised action, while transition errors operate on the arrows by changing the next state.



Action error $\epsilon_A(R)$: execute the opposite action rather than $a_m(s)$ in the current state.

Transition error $\epsilon_T(Q)$: move to the wrong next state after observing the realised action profile.

Figure 3: Automaton representation of the strategies and the two error channels. State labels are prescribed actions. Arrow labels are the counterpart's realised action in the previous period. Action errors are linked to rationality R; transition errors are linked to intelligence Q. The notation xS–yT follows the convention that x is the number of states and y is the number of transitions in excess of the number of states; "a" denotes an absorbing state.

4.2 Two ability-linked error probabilities

Each player i is characterised by a pair (R_i, Q_i), where $R_i \in [0, 1]$ is the fGARP-based rationality score and $Q_i \in [0, 1]$ is a normalised measure of intelligence. A player implementing strategy m in state s first forms the intended action $a_m(s)$. The realised action $\tilde{a}_i$ is then

$$\Pr(\tilde{a}_i = a_m(s) \mid s, m, R_i) = 1 - \epsilon_A(R_i),$$

The probability $\epsilon_A(R_i)$ is the action-error probability. The central restriction is

$$\epsilon_A'(R_i) < 0,$$

so higher rationality reduces the probability that a player fails to execute the action prescribed by the current state. This captures the interpretation of rationality as consistency in implementation: a highly rational subject is not necessarily choosing a different rule, but is less likely to deviate from the rule she is currently trying to follow.

After the realised action profile $(\tilde{a}_i, \tilde{a}_j)$ is observed, the automaton prescribes the next state $\phi_m(s, \tilde{a}_i, \tilde{a}_j)$. The realised next state $\tilde{s}'$ satisfies

$$\Pr\left(\tilde{s}' = \phi_m(s, \tilde{a}_i, \tilde{a}_j) \mid s, m, Q_i\right) = 1 - \epsilon_T(Q_i),$$

and, when there is more than one state, the remaining probability is assigned to an incorrect state. The probability $\epsilon_T(Q_i)$ is the transition-error probability. The central restriction is

$$\epsilon_T'(Q_i) < 0,$$

so higher intelligence reduces the probability that a player loses track of the automaton's state or updates it incorrectly. Transition errors have no behavioural content for one-state strategies such as AD and AC; they matter only for contingent strategies such as GT and TFT.

This separation is the key departure from a single cognitive-ability interpretation of errors. Both R_i and Q_i may be correlated, but they enter the model through different operations: R_i governs the reliability of actions conditional on a state, while Q_i governs the reliability of state updating conditional on the realised history.

4.3 Payoffs with errors

Let $V_{x,y}(m, n)$ denote the expected normalised payoff to a player of type $x = (R, Q)$ who uses strategy m against a player of type $y = (R', Q')$ who uses strategy n. Formally,

$$V_{x,y}(m, n) = (1 - \delta)\, \mathbb{E}_{x,y,m,n}\left[\sum_{t=1}^{\infty} \delta^{t-1} u(\tilde{a}_t^i, \tilde{a}_t^j)\right],$$

where the expectation is taken over both players' action and transition errors. The normalising factor $(1-\delta)$ is immaterial for best responses but is convenient because it puts payoffs on the same scale as per-period payoffs.

For a group g with distribution of types F_g and current population distribution of strategies $\mu \in \Delta(M)$, the expected payoff to type x from strategy m is

$$U_g(m, \mu; x) = \sum_{n \in M} \mu(n) \int V_{x,y}(m, n)\, dF_g(y).$$

For aggregate comparisons across rationality groups it is useful to work with the group-average payoff

$$\bar{U}_g(m, \mu) = \int U_g(m, \mu; x)\, dF_g(x).$$

The high-rationality group differs from the low-rationality group primarily through a lower average action-error probability,

$$\bar{\epsilon}_A^H = \int \epsilon_A(R)\, dF_H(R, Q) \;<\; \bar{\epsilon}_A^L = \int \epsilon_A(R)\, dF_L(R, Q),$$

whereas differences in transition errors are governed by the distribution of intelligence within each group,

$$\bar{\epsilon}_T^g = \int \epsilon_T(Q)\, dF_g(R, Q).$$

Thus the design isolates a rationality channel even though intelligence is also measured and can be controlled for empirically.

4.4 Adjustment and basins of attraction

Following the best-response adjustment model, suppose that in a short time interval dt a fraction λdt of the population revises its strategy while the remaining fraction $1 - \lambda dt$ keeps its current strategy. Revising players best respond to the current population distribution. Let

$$BR_g(\mu) = \left\{ \tau \in \Delta(M) : \bar{U}_g(\tau, \mu) \geq \bar{U}_g(m, \mu) \text{ for all } m \in M \right\}.$$

Then, for every strategy $m \in M$,

$$\mu_g(t + dt, m) = (1 - \lambda dt)\mu_g(t, m) + \lambda dt\, \tau(m), \qquad \tau \in BR_g(\mu_g(t)).$$

Taking the limit as $dt \to 0$ gives the differential inclusion

$$\dot{\mu}_g(t, m) = \lambda \Big(\tau(m) - \mu_g(t, m)\Big), \qquad \tau \in BR_g(\mu_g(t)).$$

This is the direct analogue of the best-response dynamics in Proto et al. (2022), with the payoff matrix now endogenously determined by both action errors and transition errors.

For any strategy m, define its basin of attraction in group g as the set of initial population states that converge to m under these dynamics:

$$\mathcal{B}_g(m) = \left\{\mu_0 \in \Delta(M) : \lim_{t\to\infty} \mu_g(t, m) = 1 \;\middle|\; \mu_g(0) = \mu_0\right\}.$$

If initial conditions are drawn from a distribution Ψ on the simplex, the size of the basin is

$$b_g(m) = \int_{\Delta(M)} \mathbf{1}\{\mu_0 \in \mathcal{B}_g(m)\}\, d\Psi(\mu_0).$$

The basin $b_g(m)$ is a summary measure of how robust strategy m is to uncertainty about the initial composition of the population. A larger basin means that a wider set of initial beliefs or early experiences lead the group toward that strategy.

4.5 Predictions for rationality and intelligence

The framework generates three predictions that organise the empirical analysis.

First, rationality should predict action errors. Since $\epsilon_A'(R) < 0$, high-rationality subjects should make fewer deviations from the action prescribed by their current strategic state. In the language of the estimation, the fGARP-based CCEI should be negatively associated with $\hat{\epsilon}_A$ even after controlling for Raven scores and other individual characteristics.

Second, intelligence should predict transition errors. Since $\epsilon_T'(Q) < 0$, higher Raven scores should be associated with fewer failures to update the strategic state correctly. This channel is especially important for multi-state strategies such as GT and TFT.

Third, the two error channels should affect strategy dynamics through different mechanisms. Higher transition-error rates make contingent cooperation harder to sustain because players are more likely to move to the wrong punishment or cooperation state after observing the previous-period action profile. Higher action-error rates undermine cooperation more directly by injecting unintended defections into the realised history, even when the player is in a cooperative state and intends to cooperate. Since such unintended actions trigger responses by others, high action-error environments should expand the basin of harsher strategies, particularly AD and GT, and shrink the basin of forgiving cooperative strategies such as TFT. Conversely, high rationality lowers action errors, making cooperative contingent strategies more viable and allowing learning to move subjects away from simple harsh rules.

These predictions clarify the distinction between rationality and intelligence. Intelligence governs the ability to track where one is in a strategy; rationality governs the discipline with which one executes the action prescribed at that point. The model therefore predicts the empirical pattern documented in Section 4: high-rationality groups can increasingly sustain cooperation because their members become more reliable implementers of their intended actions. The model also explains why rationality and intelligence should not be treated as interchangeable. Even when the two measures are positively correlated, they enter the dynamics through different components of the automaton.

5 Strategy Identification and Error Decomposition

To identify the strategies adopted by subjects and to quantify the consistency with which these strategies are implemented, we adopt a maximum likelihood (ML) estimation framework. Building on the theoretical framework in Section 4 our task will be to recover the action error probability ϵ_A and the transition error probability ϵ_T for each participant. Recall that ϵ_T reflects the failure to update the state of the strategy automaton correctly, given the realised action profile.

Transition errors therefore capture failures in tracking or updating the strategic state and are therefore naturally associated with limitations in cognitive processing and working memory. Similarly, recall that ϵ_A reflects the selection of an action that differs from the one prescribed by the automaton's current state. These errors capture momentary lapses, inattention, or lack of care, rather than misunderstanding of the strategic rule itself. Importantly, action errors are the only source of stochasticity for single-state strategies such as Always Defect or Always Cooperate, making their inclusion necessary for a coherent likelihood-based estimation.

5.1 Likelihood Construction

A key challenge in identifying strategies from observed play is that the same action can often be rationalised by different unobserved cognitive paths. For example, a defection may reflect either an action error or a deliberate transition to a punishment state following the partner's behaviour. To address this, the model considers all possible sequences of latent states and errors that are consistent with a given candidate strategy.

The likelihood of a strategy is defined as the sum of the probabilities of all admissible error histories that could have generated the observed sequence of choices. This is computed using a recursive algorithm that progresses through the history of play, tracking the set of feasible state transitions and associated error realizations at each step. The resulting likelihood function jointly depends on the strategy and the two error probabilities (ϵ_T, ϵ_A).

5.2 Estimation Procedure

To balance the need for sufficient observations for identification against the possibility that subjects adjust strategies over time, estimation is conducted on blocks of five consecutive supergames, following Proto et al. (2022). For each block, the procedure identifies the strategy and error probabilities that maximise the likelihood of the observed action sequence.

This approach yields individual-level estimates of both strategy choice and implementation consistency, providing a disciplined framework for decomposing observed behaviour into strategic intent and execution errors. In the analysis that follows, we exploit this structure to examine

how individual characteristics, most notably rationality, are related to distinct forms of strategic inconsistency in repeated interactions.

Table 5 summarises the results of the maximum-likelihood estimations, reporting the most likely strategies and the estimated probabilities of transition errors and action errors for each rationality group in the first and second halves of the session. The difference columns facilitate both within-group comparisons across session halves, columns (1)-(2) for the low-rationality group and columns (3)-(4) for the high-rationality group, as well as cross-group comparisons within each half.

Table 5: Error Decomposition: Group Means and Pairwise Comparisons

	Group Means				Differences			
	Low Rationality		High Rationality		(1)–(2)	(1)–(3)	(2)–(4)	(3)–(4)
	First Half (1)	Second Half (2)	First Half (3)	Second Half (4)				
Action	0.1249	0.0845	0.0854	0.0242	0.040***	0.039**	0.060***	0.061***
Transition	0.0393	0.0375	0.0264	0.0182	0.002	0.013*	0.019**	0.008
AC	0.0955	0.0729	0.1406	0.0625	0.023	-0.045*	0.010	0.078**
AD	0.3924	0.4028	0.2240	0.1458	-0.010	0.168***	0.257***	0.078**
TfT	0.2630	0.2595	0.3073	0.3620	0.003	-0.044	-0.102***	-0.055*
GT	0.2491	0.2700	0.3281	0.4297	-0.021	-0.079**	-0.160***	-0.102***

Notes. This table reports the estimated shares of maximum-likelihood (ML) strategies and the associated error probabilities for the first and second halves of each session. Estimates are obtained using the algorithm proposed by Proto et al. (2022). Because some ML strategies are not uniquely identified, that is, their likelihood coincides with that of at least one alternative strategy, the reported strategy shares do not necessarily sum to one. Differences are based on two-sided t-tests. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Across both rationality groups, the probability of action errors (ϵ_A) declines significantly from the first to the second half of the session, indicating improved consistency in implementing a given strategy over time. Table 5 further shows that this improvement is not uniform across groups. While both low- and high-rationality subjects reduce action errors with experience, comparisons between columns (1) and (3) in the first half and columns (2) and (4) in the second half reveal that the initial gap in action-error rates across rationality groups widens over time.

While a direct comparison is not straightforward, since subjects in Proto et al. (2022) play approximately 20 supergames in each half of the session, compared to 15 in our design, it is nevertheless informative to note that the pattern observed here contrasts with their findings. In their IQ-split treatments, Proto et al. (2022) document a larger reduction in action-error

rates among low-IQ subjects over the course of the session, whereas in our data the decline in action errors is steeper for high-rationality subjects. Moreover, they report lower average error probabilities for both high- and low-IQ subjects than those observed in our experiment, suggesting differences in overall noise levels across settings.

By contrast, transition errors (ϵ_T) exhibit substantially less sensitivity to experience within the session. Transition-error rates do not differ materially across rationality groups in either the first or the second half, and the within-group decline from the first to the second half is modest for both groups. This pattern suggests that errors related to tracking and updating the strategic state are less responsive to short-run experience, consistent with the view that such errors reflect more cognitively demanding operations. The limited responsiveness of ϵ_T to experience mirrors the findings of Proto et al. (2022), who document a similar distinction when cognitive ability is proxied by IQ.

Turning to strategy composition, clear differences emerge across rationality groups. Low-rationality subjects rely more heavily on harsh strategies, most notably AD and less on co-operative or conditional strategies such as GT in both halves of the session, and on TfT in the second half, relative to high-rationality subjects (the difference between columns (1) and (3) for the first half and columns (2) and (4) for the second half). Moreover, the distribution of strategies among low-rationality subjects remains largely stable over time, indicating little adjustment in strategic choice across session halves.

By contrast, high-rationality subjects exhibit a marked reallocation of strategy use over the course of the session. In the second half, they substantially reduce their reliance on simpler strategies such as Always Defect and Always Cooperate, and instead adopt more sophisticated contingent strategies, including TfT and GT, more frequently. Importantly, this shift towards more complex strategies coincides with a pronounced decline in action-error rates, indicating improved implementation despite increased strategic complexity.

Taken together, the evidence from strategy composition and error frequencies indicates that rationality is more closely associated with a disciplined and consistent execution of strategies

rather than the ability to track where one is in a strategy. While Proto et al. (2022) show that intelligence primarily predicts errors in transition, our results suggest that rationality operates mainly through action errors, shaping how reliably a chosen strategy is implemented. The convergence in cooperative behaviour observed in our experiment is therefore driven largely by improvements in execution rather than by fundamental changes in strategic reasoning: action errors decline sharply with experience, particularly among high-rationality subjects, whereas transition errors remain comparatively stable. This distinction reinforces the view that rationality and intelligence capture different dimensions of decision-making, as emphasized in Section Section 4.

In Appendix D, we extend the analysis by exploiting the full distribution of rationality and conducting a more disaggregated examination across rationality quartiles. The quartile-level analysis yields results that are fully consistent with the main findings based on the high–low rationality split. Action errors decline over the course of the session for all quartiles, but the reduction is largest for subjects in the highest rationality quartile, leading to a widening gap over time. By contrast, transition errors display comparatively small changes across session halves and limited differentiation across quartiles. These patterns reinforce the conclusion that rationality primarily affects the consistency with which strategies are implemented, rather than the updating of states within a strategy rule.

As a robustness check on the blockwise strategy classification, Appendix E allows latent strategy use to persist across adjacent blocks using a hidden Markov model, while retaining the same within-block likelihood and error structure. The HMM preserves the main results: high-rationality environments exhibit less unconditional defection and greater use of conditional strategies, while the estimated differential decline in action errors is essentially unchanged.

6 Conclusion

This paper investigates how rationality shapes cooperation in strategic settings by linking a direct measure of rationality, based on consistency with the Generalized Axiom of Revealed Preference (GARP), to behaviour in an indefinitely repeated Prisoner's Dilemma. By grouping participants according to their measured rationality and analysing both behavioural outcomes and underlying error processes, the paper provides the first account of how decision-making consistency affects cooperation.

The results establish three main findings. First, higher rationality is associated with substantially higher levels of cooperation and payoffs. These differences are not present at the outset of play but emerge and widen over time, indicating that they are driven by learning and adaptation during repeated interaction rather than initial behavioural differences. High-rationality subjects increasingly sustain cooperation and are more likely to initiate cooperative play in later supergames, while low-rationality subjects exhibit limited adjustment.

Second, the analysis of strategies and error decomposition shows that rationality operates primarily through implementation rather than updating within a given strategy. High-rationality individuals make significantly fewer action errors, deviations from intended actions, and reduce these errors more rapidly with experience. In contrast, transition errors, which relate to failures in updating strategic states, display limited variation across rationality groups and remain relatively stable over time. This asymmetry implies that rationality mainly enhances the consistency with which strategies are executed, rather than the ability to track or update them.

Third, these findings clarify the distinction between rationality and intelligence. While both are positively associated with cooperation, they operate through different mechanisms. Consistent with the framework discussed in Section 4, intelligence is linked to transition errors and cognitive processing, whereas rationality is linked to action errors and decision-making consistency. This distinction provides empirical support for the view that rationality and intelligence capture different dimensions of behaviour, and should not be treated as interchangeable concepts.

Taken together, the evidence demonstrates that rationality plays a central role in sustaining cooperation by reducing implementation errors in repeated interactions. By directly measuring rationality and decomposing behaviour into distinct error types, the paper contributes to a more precise understanding of how cognitive constraints shape strategic behaviour. More broadly, the results suggest that policies or environments that improve decision-making consistency may enhance cooperative outcomes, even when underlying strategic understanding remains unchanged.

Appendix

A Afriat and Varian Quartile Concordance

Table A1: Correlation Matrix for Rationality Quartile Assignments (Exact Sample, $N = 196$)

Measure	CCEI (Afriat)	Varian (Mean)	Varian (SSQ)
CCEI (Afriat)	1.0000		
Varian (Mean)	0.9484	1.0000	
Varian (SSQ)	0.9743	0.9649	1.0000

Note: Correlations are calculated using the subsample of subjects ($N = 196$) where Varian indices were solved exactly using the procedure from Halevy et al. (2018). All measures refer to quartile rankings derived under the standard GARP axiom to ensure compatibility across the Afriat and Varian frameworks.

Table A2: Counterfactual Comparison of Afriat and Varian Quartile Assignments

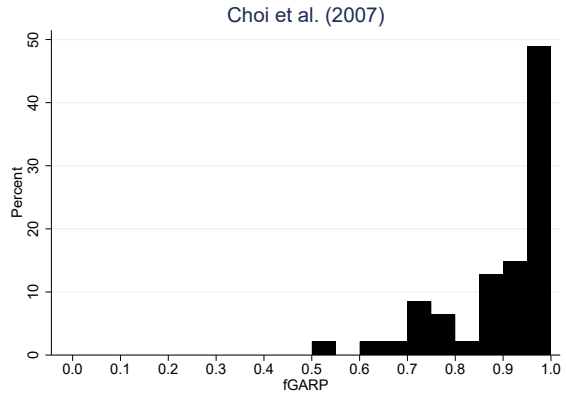
Measure	Agreement (%)	Kappa (κ)	Quartile Moves (%)	Interpretation
Varian (Mean)	93.88%	0.9017	6.1%	Very Stable
Varian (SSQ)	96.94%	0.9510	3.1%	Even More Stable

Note: To replicate the experimental grouping protocol, counterfactual quartiles are calculated at the session level. Specifically, for each session, subjects are ranked by their respective inconsistency indices and split into quartiles, mirroring the original assignment process based on the Choi task CCEI. The Kappa (κ) statistic measures the agreement between the Afriat-based and Varian-based assignments while accounting for the agreement that could occur by chance; values above 0.90 indicate near-perfect concordance. Comparisons use the GARP axiom to ensure compatibility with the Halevy et al. (2018) estimation code. "Moves" refers to the percentage of subjects shifting by exactly one quartile; no subjects shifted by more than one quartile.

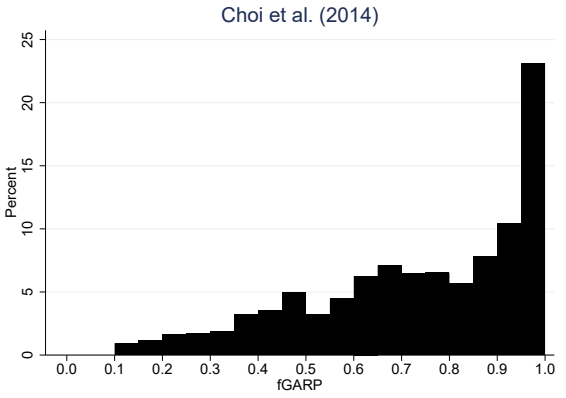
B Distribution of fGARP Across Studies

Table A3: Distribution of FGARP by study

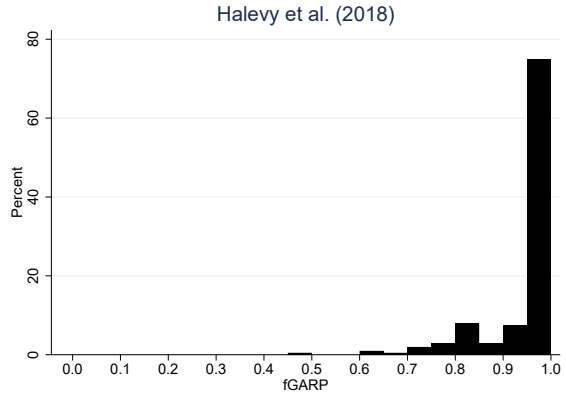
Study	Mean	SD	p10	p25	p50	p75	p90	p95	*N*	# Tasks
Choi et al. (2007)	0.898	0.116	0.709	0.854	0.947	0.991	0.999	0.999	47	50
Choi et al. (2014)	0.733	0.228	0.393	0.584	0.775	0.943	0.981	0.991	1,182	25
Halevy et al. (2018)	0.952	0.083	0.827	0.950	0.987	1.000	1.000	1.000	203	22
This study's pilot	0.923	0.089	0.814	0.885	0.961	0.986	0.994	0.999	32	25
This study	0.832	0.191	0.545	0.736	0.913	0.975	0.994	1.000	256	25



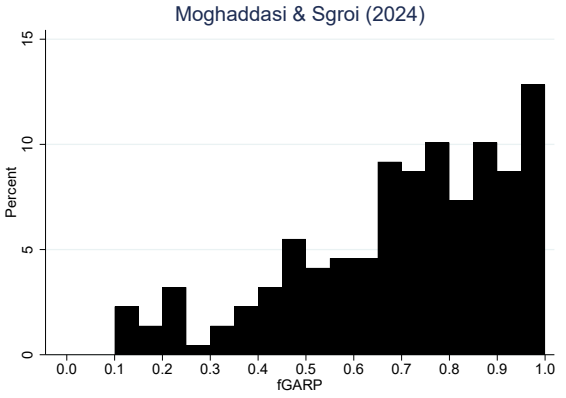
(a) Choi et al. (2007)



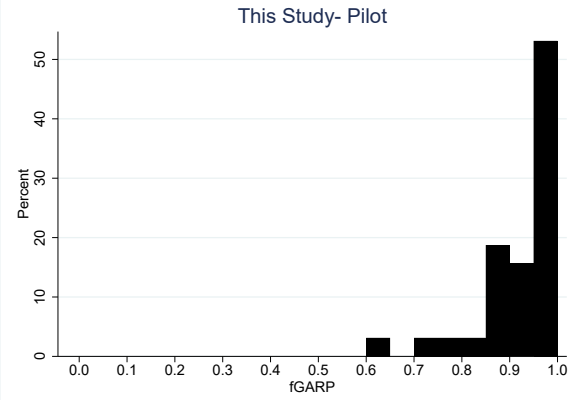
(b) Choi et al. (2014)



(c) Halevy et al. (2018)



(d) Moghaddasi & Sgroi (2024)



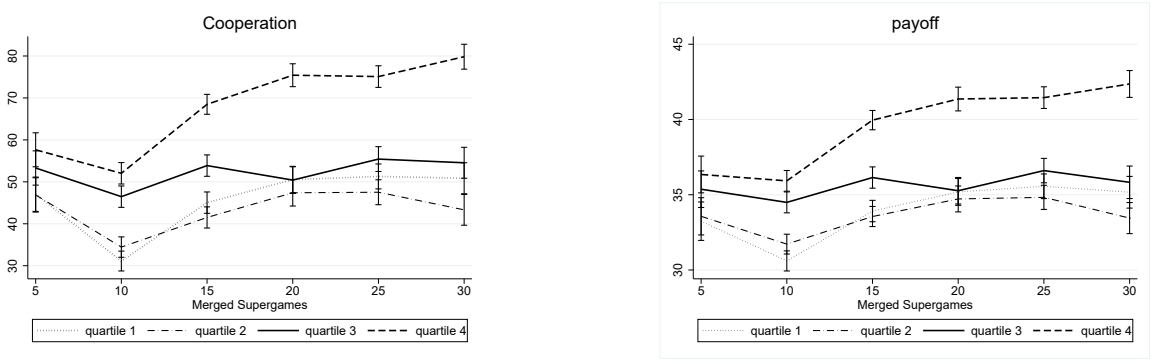
(e) Present study- pilot

Figure A1: Distributions of fGARP across studies

Notes: Each panel reports the percentage distribution of individual fGARP scores using common bins and a common horizontal scale from zero to one. Panel (a) uses the 47 participants in the symmetric treatment of Choi et al. (2007), drawn from a laboratory sample recruited from undergraduate classes and staff at the University of California, Berkeley. Panel (b) uses 1,182 adult members of the Dutch CentERpanel studied by Choi et al. (2014); the panel was designed to be representative of the Dutch-speaking adult population of the Netherlands. Panel (c) uses the 203 participants in Halevy et al. (2018), who were recruited at the University of British Columbia and were primarily undergraduate students. Panel (d) uses the 218 participants in Moghaddasi Kelishomi and Sgroi (2024), who were recruited from living unrelated kidney donors in Iran. Panel (e) uses the 32 university-student participants in the present study's pilot. Higher values indicate greater consistency with fGARP.

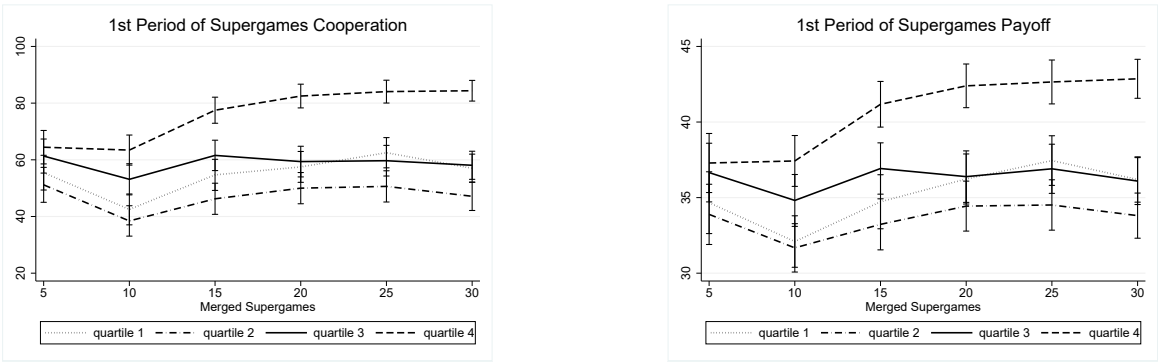
C Cooperation and Payoff by Rationality Quartiles

Figures A2 and A3 show cooperation and payoffs separately for each rationality quartile. Quartile 4 consistently stands out from the lower three quartiles, supporting the high- versus low-rationality split used in the main analysis.



(a) Cooperation (b) Payoff

Figure A2: Average cooperation (left) and Payoff (right) computed over observations in successive blocks of five supergames, shown separately for sessions in each rationality quartile. Bands represent 95 percent confidence intervals.



(a) Cooperation (b) Payoff

Figure A3: Average cooperation (left) and Payoff (right) in the first rounds (periods) of supergames, computed over observations in successive blocks of five supergames, shown separately for sessions in each rationality quartile. Bands represent 95 percent confidence intervals.

D Quartile-Level Error Decomposition

Table A4 presents a quartile-level decomposition of action errors across session halves. In the first half of the session, action-error rates decline monotonically with rationality, with the largest and most statistically significant differences arising between the lowest and highest quartiles. Importantly, this gradient is nonlinear: the separation between the top quartile and the remaining

subjects is sharper than differences among the lower quartiles, consistent with the distributional properties of the rationality measure.

Across all quartiles, action errors decline significantly from the first to the second half of the session, indicating that experience improves implementation consistency. However, the magnitude of this improvement differs markedly across quartiles. Subjects in the highest rationality quartile exhibit the largest absolute reduction in action errors despite starting from a lower baseline, resulting in a widening gap over time. This pattern confirms that experience interacts with rationality by operating at different speeds: while all subjects improve with repeated interaction, high-rationality individuals reduce implementation errors more rapidly.

These findings corroborate the main results obtained from the binary rationality split and demonstrate that the learning dynamics observed there are not driven by aggregation across heterogeneous types, but instead reflect a robust monotonic pattern concentrated at the upper end of the rationality distribution.

Table A4: Action Errors by Rationality Quartile and Session Half

Outcome	Comparison	Group A	Mean(A)	Group B	Mean(B)	Diff(A−B)	p-value
action	FH vs FH	q1-FH	0.148763	q2-FH	0.118164	0.030599	0.0463
action	FH vs FH	q1-FH	0.148763	q3-FH	0.107747	0.041016	0.0059
action	FH vs FH	q1-FH	0.148763	q4-FH	0.085417	0.063346	0.0000
action	FH vs FH	q2-FH	0.118164	q3-FH	0.107747	0.010417	0.4674
action	FH vs FH	q2-FH	0.118164	q4-FH	0.085417	0.032747	0.0228
action	FH vs FH	q3-FH	0.107747	q4-FH	0.085417	0.022331	0.1060
action	SH vs SH	q1-SH	0.093815	q2-SH	0.083529	0.010286	0.4865
action	SH vs SH	q1-SH	0.093815	q3-SH	0.076042	0.017773	0.2045
action	SH vs SH	q1-SH	0.093815	q4-SH	0.024154	0.069661	0.0000
action	SH vs SH	q2-SH	0.083529	q3-SH	0.076042	0.007487	0.5881
action	SH vs SH	q2-SH	0.083529	q4-SH	0.024154	0.059375	0.0000
action	SH vs SH	q3-SH	0.076042	q4-SH	0.024154	0.051888	0.0000
action	FH vs SH	q1-FH	0.148763	q1-SH	0.093815	0.054948	0.0004
action	FH vs SH	q2-FH	0.118164	q2-SH	0.083529	0.034635	0.0192
action	FH vs SH	q3-FH	0.107747	q3-SH	0.076042	0.031706	0.0183
action	FH vs SH	q4-FH	0.085417	q4-SH	0.024154	0.061263	0.0000

Notes. Diff(A−B) = Mean(A) − Mean(B). p-values are from two-sided t-tests. FH and SH denote the first and second halves of the session.

Table A5 reports the corresponding quartile-level analysis for transition errors. In contrast to action errors, transition-error rates exhibit relatively weak differentiation across rationality quar-

tiles and limited responsiveness to experience within the session. While the highest rationality quartile displays lower transition-error rates than the lowest quartile in both halves of the session, intermediate quartiles are not systematically distinguishable, and within-quartile changes over time are small and statistically insignificant.

Moreover, the reduction in transition errors from the first to the second half of the session is modest across all quartiles, with no evidence of accelerated improvement among higher-rationality subjects. This stability contrasts sharply with the pronounced decline observed for action errors and suggests that errors related to tracking and updating the strategic state are less amenable to short-run learning.

Taken together, the quartile analysis reinforces the asymmetric role of rationality across error types identified in the main text. Rationality is strongly associated with the dynamics of action errors, but plays a more limited role in shaping the evolution of transition errors, consistent with the interpretation that the latter reflect more cognitively demanding operations.

Table A5: Transition Errors by Rationality Quartile and Session Half

Outcome	Comparison	Group A	Mean(A)	Group B	Mean(B)	Diff(A−B)	p-value
transition	FH vs FH	q1-FH	0.050977	q2-FH	0.029297	0.021680	0.0241
transition	FH vs FH	q1-FH	0.050977	q3-FH	0.037500	0.013477	0.2014
transition	FH vs FH	q1-FH	0.050977	q4-FH	0.026432	0.024544	0.0094
transition	FH vs FH	q2-FH	0.029297	q3-FH	0.037500	-0.008203	0.3799
transition	FH vs FH	q2-FH	0.029297	q4-FH	0.026432	0.002865	0.7216
transition	FH vs FH	q3-FH	0.037500	q4-FH	0.026432	0.011068	0.2273
transition	SH vs SH	q1-SH	0.043945	q2-SH	0.037240	0.006706	0.5545
transition	SH vs SH	q1-SH	0.043945	q3-SH	0.031185	0.012760	0.2420
transition	SH vs SH	q1-SH	0.043945	q4-SH	0.018164	0.025781	0.0077
transition	SH vs SH	q2-SH	0.037240	q3-SH	0.031185	0.006055	0.5649
transition	SH vs SH	q2-SH	0.037240	q4-SH	0.018164	0.019076	0.0387
transition	SH vs SH	q3-SH	0.031185	q4-SH	0.018164	0.013021	0.1326
transition	FH vs SH	q1-FH	0.050977	q1-SH	0.043945	0.007031	0.5315
transition	FH vs SH	q2-FH	0.029297	q2-SH	0.037240	-0.007943	0.4134
transition	FH vs SH	q3-FH	0.037500	q3-SH	0.031185	0.006315	0.5350
transition	FH vs SH	q4-FH	0.026432	q4-SH	0.018164	0.008268	0.2654

Notes. Diff(A−B) = Mean(A) − Mean(B). p-values are from two-sided t-tests. FH and SH denote the first and second halves of the session.

E Temporal Persistence and Hidden-Markov Robustness

The strategy estimates reported in the main text follow the recursive procedure of Proto et al. (2022). For each subject, the thirty supergames are divided into six consecutive blocks of five supergames. Within each block, the likelihood of strategies: Always Defect (AD), Always Cooperate (AC), Tit-for-Tat (TfT), and Grim Trigger (GT) is evaluated after maximising over the corresponding action- and transition-error history probabilities for each sequence of observed actions. Blocks are classified independently, and ties between maximum-likelihood strategies are handled using the allocation convention described in the main text.

In this part, we ask whether the main strategy and error results are robust to allowing strategies to be persistent across adjacent blocks. We do so using a hidden Markov model (HMM). The hidden state is the strategy used by a subject in a given five-supergame block, while the observed data are the strategy-specific likelihoods produced by the original blockwise estimator. The HMM therefore does not replace the within-block error-estimation procedure. Instead, it adds a temporal smoothing layer that uses information from neighbouring blocks when assigning posterior probabilities to strategies.

Before estimating the HMM, we first check whether persistence is already visible in the independent blockwise classifications. This provides a direct diagnostic for whether adding a dynamic classification layer is empirically motivated.

E.1 Persistence in the blockwise estimates

Before imposing a dynamic model, we examine whether the independently estimated classifications already display continuity across adjacent blocks. Let w_{ibm} denote the original classification weight assigned to strategy m for subject i in block b. We measure fractional adjacent-block persistence by

$$\widehat{P}^{\text{same}} = \frac{1}{N(B-1)} \sum_{i=1}^{N} \sum_{b=1}^{B-1} \sum_{m} w_{ibm} w_{i,b+1,m}. \tag{1}$$

The benchmark under serial independence is

$$\widehat{P}^{\text{chance}} = \sum_{m} \bar{w}_m^2, \qquad \bar{w}_m = \frac{1}{NB} \sum_{i,b} w_{ibm}. \tag{2}$$

Panel A of Table A6 shows substantial continuity in the original blockwise estimates. Observed adjacent-block persistence is 0.550, compared with 0.288 under independent draws from the marginal strategy shares. The corresponding chance-adjusted persistence measure is 0.368. Moreover, among the 729 adjacent block pairs for which both blocks have a unique maximum-likelihood classification, the same strategy is selected in 65.2% of cases. Thus, persistence is already visible in the independently estimated classifications and is not solely an implication of the HMM.

E.2 Hidden-Markov specification

The HMM is added as a second-stage extension of the original blockwise estimator. It does not replace the recursive treatment of action and transition errors within a block. Instead, it uses the complete set of strategy-specific likelihoods produced by that estimator and introduces a Markov process governing strategy use across adjacent blocks.

Let Y_{ib} denote the observed sequence of own and partner actions for subject i in block b, and let

$$m \in \mathcal{M} = \{\text{AD}, \text{AC}, \text{TfT}, \text{GT}\}$$

index the four candidate strategies. For each subject–block–strategy combination, the original procedure computes

$$L^*_{ibm} = \max_{\epsilon^A, \epsilon^T} \Pr\left(Y_{ib} \mid m, \epsilon^A, \epsilon^T\right), \tag{3}$$

where the probability of the observed action history sums over all latent sequences of action errors, transition errors, and internal automaton states that are consistent with that history. The corresponding maximisers,

$$\widehat{\epsilon}^A_{ibm} \quad \text{and} \quad \widehat{\epsilon}^T_{ibm},$$

are retained for the subsequent construction of posterior-weighted error rates. Transition errors are not applicable to the one-state strategies AD and AC.

The stored L^*_{ibm} values are likelihood levels rather than log likelihoods. We therefore use them directly as state-dependent emission scores. For numerical convenience, they are normalised within each subject–block:

$$E_{ibm} = \frac{L^*_{ibm}}{\max_{r \in \mathcal{M}} L^*_{ibr}}. \tag{4}$$

This normalisation leaves all likelihood ratios unchanged. It therefore has no effect on posterior strategy probabilities or on the estimated transition parameters, because it multiplies the likelihood of every possible strategy path for a given subject by the same sequence of block-specific constants.

Let

$$S_{ib} \in \mathcal{M}$$

denote the latent strategy used by subject i in block b. The HMM initial-state probabilities and transition probabilities are

$$\pi_m = \Pr(S_{i1} = m)$$

and

$$A_{sr} = \Pr(S_{ib} = r \mid S_{i,b-1} = s),$$

respectively. For a subject with $B = 6$ blocks, the sequence likelihood is

$$\mathcal{L}_i(\boldsymbol{\pi}, A) = \sum_{s_1,\ldots,s_B} \pi_{s_1} E_{i1 s_1} \prod_{b=2}^{B} A_{s_{b-1} s_b} E_{ib s_b}. \tag{5}$$

The sample log likelihood is

$$\ell(\boldsymbol{\pi}, A) = \sum_{i=1}^{N} \log \mathcal{L}_i(\boldsymbol{\pi}, A). \tag{6}$$

The summation in equation (5) integrates over all 4^6 possible strategy paths, but is evaluated efficiently using the forward algorithm.

We report three specifications to separate three issues: whether soft posterior classification alone

matters, whether temporal persistence improves classification, and whether the parsimonious persistence restriction is too strong. All three specifications use the same emission scores, namely the subject–block–strategy profile likelihoods from the original estimator. They differ only in how they restrict the latent transition matrix A, which governs the evolution of strategies across adjacent blocks.

The independent mixture provides the no-persistence benchmark. It estimates the marginal frequency of each strategy but assumes that a subject's strategy in one block contains no information about the strategy used in the next block. The sticky mixture adds one persistence parameter and is our preferred specification. It allows the previous block's strategy to persist while still allowing switches to be directed by the estimated population distribution of strategies. The unrestricted HMM is included only as a flexibility check. It estimates every row of the transition matrix separately and therefore allows, for example, switches out of AD to differ from switches out of TfT.

Independent latent mixture. The first specification permits unequal strategy frequencies but assumes no dependence across blocks. Let

$$\boldsymbol{q} = (q_{\mathrm{AD}}, q_{\mathrm{AC}}, q_{\mathrm{TfT}}, q_{\mathrm{GT}})$$

denote the estimated population distribution of strategies. The initial and transition probabilities are

$$\boldsymbol{\pi} = \boldsymbol{q}, \qquad A_{sr} = q_r \quad \text{for all } s, r. \tag{7}$$

Thus, every row of A equals $\boldsymbol{q}'$, and the latent strategy in each block is drawn independently from the same non-uniform distribution. This model has three free parameters.

Sticky latent mixture. Our preferred specification adds a single persistence parameter:

$$A(\rho, \boldsymbol{q}) = \rho I + (1 - \rho)\mathbf{1}\boldsymbol{q}', \qquad 0 \le \rho \le 1, \tag{8}$$

with

$$\boldsymbol{\pi} = \boldsymbol{q}.$$

With probability ρ, the subject retains the strategy used in the preceding block.[6] With probability $1 - \rho$, the strategy is redrawn from $\boldsymbol{q}$. Hence,

$$A_{ss} = \rho + (1 - \rho)q_s \tag{9}$$

and, for $r \neq s$,

$$A_{sr} = (1 - \rho)q_r. \tag{10}$$

Importantly, $\rho = 0$ represents serial independence without imposing equal frequencies of the four strategies. The sticky specification has four free parameters: three elements of $\boldsymbol{q}$ and ρ.

Unrestricted pooled HMM. The final specification estimates a separate initial distribution $\boldsymbol{\pi}$ and a full 4×4 transition matrix:

$$A_{sr} \geq 0, \qquad \sum_{r \in \mathcal{M}} A_{sr} = 1 \quad \text{for each } s. \tag{11}$$

This model allows persistence and switching destinations to differ freely across origin strategies. It has fifteen free parameters: three initial-state parameters and twelve transition parameters. We include it to check whether the parsimonious sticky restriction is too strong, but we do not use it for the substantive comparisons because, as shown below, it improves fit only slightly relative to the sticky model while adding many parameters.

The independent mixture is estimated by expectation maximisation. The sticky model is estimated by direct maximum likelihood, using multiple starting values for both ρ and $\boldsymbol{q}$. The unrestricted model is estimated using the Baum–Welch algorithm with multiple starting val-

[6]For illustration, suppose there are four strategies and $\boldsymbol{q} = (0.30, 0.10, 0.30, 0.30)$. If $\rho = 0$, each row of the transition matrix equals $\boldsymbol{q}$, so the strategy in the next block is independent of the current strategy. If $\rho = 0.80$, then a subject currently in AD remains in AD with probability $0.80 + 0.20 \times 0.30 = 0.86$, and switches to AC, TfT, or GT with probabilities $0.20 \times 0.10 = 0.02$, $0.20 \times 0.30 = 0.06$, and $0.20 \times 0.30 = 0.06$, respectively. The same rule applies to every row: the diagonal element receives the persistence component ρ, while off-diagonal probabilities are proportional to the estimated population shares in $\boldsymbol{q}$.

ues. In each specification, the forward–backward algorithm produces the smoothed posterior probabilities

$$\gamma_{ibm} = \Pr(S_{ib} = m \mid Y_{i1}, \ldots, Y_{iB}) . \tag{12}$$

These probabilities use information from both preceding and subsequent blocks and form the basis of the HMM strategy shares and posterior-weighted error estimates reported below. We additionally calculate the most likely complete strategy path using the Viterbi algorithm, but use it only for switching diagnostics rather than for the substantive outcome comparisons.

The resulting estimator is a profile-likelihood HMM rather than a fully joint structural HMM: the action- and transition-error probabilities are estimated separately for each subject–block–strategy combination before the Markov process is estimated. This construction deliberately preserves the within-block error model used in the main analysis and isolates the effect of adding temporal information to strategy classification.

Table A6: Persistence and Hidden-Markov Model Diagnostics

Panel A: Persistence in the independent block estimates			
Fractional adjacent-block persistence		0.550	
Persistence under independent marginal draws		0.288	
Chance-adjusted persistence		0.368	
Same strategy when both classifications are unique		0.652	
Adjacent pairs with two unique classifications		729	
Panel B: HMM model comparison			
	Independent mixture	Sticky mixture	Unrestricted HMM
Log likelihood	-2584.739	-2185.709	-2180.971
Parameters	3	4	15
AIC	5175.5	4379.4	4391.9
Panel C: Sticky-HMM diagnostics			
Persistence parameter, $\widehat{\rho}$		0.9596	
Expected switches per subject		0.140	
Viterbi switches per subject		0.043	
Strongly resolved TfT–GT ties		0.125	

Notes: The sample contains 256 subjects, six blocks per subject, and 1,280 adjacent-block pairs. Panel A is calculated directly from the original blockwise strategy-allocation variables. Chance-adjusted persistence is $(\widehat{P}^{\text{same}} - \widehat{P}^{\text{chance}})/(1 - \widehat{P}^{\text{chance}})$. A TfT–GT tie is classified as strongly resolved when the larger posterior probability, conditional on posterior mass on TfT and GT, is at least 0.75.

The sticky model provides a large improvement in fit relative to the independent mixture while adding only one parameter. The unrestricted model improves the log likelihood by only 4.74

despite adding eleven parameters, and AIC favours the sticky specification. The estimated persistence parameter is

$$\widehat{\rho} = 0.9596,$$

implying highly persistent latent strategy paths.

In the sticky specification, the estimated population distribution is

$$\widehat{\boldsymbol{q}} = (0.334, 0.041, 0.311, 0.314),$$

for AD, AC, TfT, and GT, respectively. Hence, a subject currently in AD is estimated to remain in AD with probability

$$0.9596 + (1 - 0.9596) \times 0.334 = 0.973,$$

while the probabilities of switching to AC, TfT, and GT are approximately 0.002, 0.013, and 0.013. This illustrates that the sticky model captures high persistence while still allowing switches to be directed by the estimated population distribution of strategies.

Temporal information nevertheless resolves only a limited proportion of the principal TfT–GT ambiguity: 12.5% of TfT–GT ties are strongly resolved. We therefore treat the HMM as a temporal-smoothing robustness exercise rather than as a definitive separation of observationally equivalent strategies. For this reason, the outcome analysis emphasises the combined posterior share of conditional strategies, TfT plus GT.

E.2.1 HMM strategy shares and implementation errors

Let

$$\gamma_{ibm} = \Pr(S_{ib} = m \mid Y_{i1}, \ldots, Y_{i6})$$

denote the smoothed posterior probability assigned to strategy m. HMM strategy shares are averages of these posterior probabilities. Posterior-weighted action errors are calculated as

$$\widehat{\epsilon}^{A,HMM}_{ib} = \sum_{m} \gamma_{ibm} \widehat{\epsilon}^{A}_{ibm}.$$

The HMM transition-error measure is defined analogously for the two-state strategies:

$$\widehat{\epsilon}^{T,HMM}_{ib} = \gamma_{ib,\mathrm{TfT}} \widehat{\epsilon}^{T}_{ib,\mathrm{TfT}} + \gamma_{ib,\mathrm{GT}} \widehat{\epsilon}^{T}_{ib,\mathrm{GT}}.$$

Thus, this measure is the unconditional posterior-weighted contribution of transition errors, with AD and AC contributing zero because transition errors are not defined for one-state strategies.

Table A7 mirrors the principal comparisons in Table 5. The first four columns report subject–block means. Statistical comparisons are obtained from

$$y_{ib} = \alpha + \beta_H H_i + \beta_D D_b + \beta_{HD}(H_i \times D_b) + \delta_s + u_{ib}, \tag{13}$$

where H_i identifies the high-rationality environment, D_b identifies the second half, and δ_s denotes session fixed effects. Standard errors are clustered at the eight-subject matching-group level.

Table A7: Sticky-HMM Strategy Shares and Implementation Errors

	Low rationality		High rationality		High–low,	Low:	Difference
Outcome	First half	Second half	First half	Second half	first half	second–first	in changes
AD share	0.387	0.387	0.187	0.140	−0.200***	−0.000	−0.047*
					(0.060)	(0.014)	(0.025)
TfT + GT share	0.577	0.577	0.754	0.794	0.177***	−0.000	0.041*
					(0.064)	(0.014)	(0.020)
Action error	0.160	0.105	0.105	0.030	−0.054**	−0.055***	−0.020
					(0.023)	(0.011)	(0.020)
Transition error	0.053	0.042	0.045	0.025	−0.007	−0.011	−0.010
					(0.009)	(0.007)	(0.011)

Notes: The first four columns report unadjusted subject–block means. The remaining columns report coefficients from equation (13). "High–low, first half" is $\widehat{\beta}_H$; "Low: second–first" is $\widehat{\beta}_D$; and "Difference in changes" is $\widehat{\beta}_{HD}$. All regressions include session fixed effects. Standard errors, in parentheses, are clustered at the eight-subject matching-group level, defined by session and rationality quartile (32 clusters). Low rationality comprises quartiles 1–3, while high rationality denotes quartile 4. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

The HMM preserves and sharpens the principal difference in strategy composition. In the first half, the posterior AD share is 20.0 percentage points lower in the high-rationality environment, while the combined posterior share of TfT and GT is 17.7 percentage points higher. Both differences are statistically significant at the 1% level.

The combined TfT–GT share is essentially unchanged between halves in the low-rationality environment. In contrast, it increases by 4.1 percentage points in the high-rationality environment:

$$0.794 - 0.754 = 0.041,$$

which is statistically significant ($p = 0.012$). Consequently, the high–low difference in conditional strategy use rises from 0.177 in the first half to 0.218 in the second half; the second-half difference is significant at the 1% level ($p = 0.003$). The corresponding difference between the two changes is positive and marginally significant ($p = 0.055$).

HMM-weighted action errors are 5.4 percentage points lower in the high-rationality environment in the first half. Action errors decline by 5.5 percentage points in the low-rationality environment, but the additional decline in the high-rationality environment is not statistically significant. Importantly, the HMM difference-in-differences estimate, -0.020, is almost identical to the corresponding estimate from the original blockwise procedure, -0.0208.

Neither the initial high–low difference nor the differential change in transition errors is statistically significant. AC is omitted from Table A7: its overall posterior share is small (approximately 4.3%), and neither its initial difference nor its differential change is statistically significant.

Overall, the HMM confirms that the main difference across rationality environments lies in strategy composition. High-rationality environments are characterised by less unconditional defection and greater use of conditional strategies, with conditional strategy use increasing further over the session. Temporal smoothing changes the precise allocation among the cooperative strategies but leaves the principal action-error result essentially unchanged.

F Experiment Instructions

Welcome to the Experiment!

Thank you everyone for coming to our experiment today.

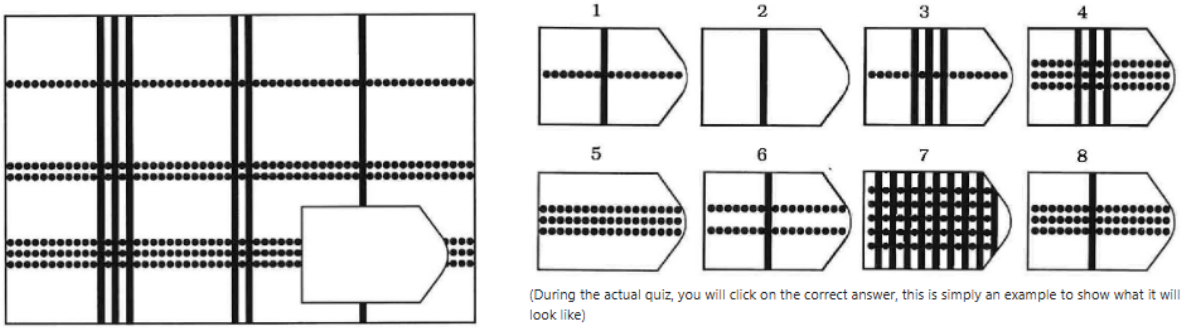
This experiment consists of 8 tasks, all of which will be completed in this session. The session will last approximately 90 minutes. You will receive your payment after completing all of the tasks.

Any questions? If you have any questions during the experiment, please raise your hand and we will come to help you. Please remain silent throughout the session as otherwise we will be forced to terminate the exercise and you will only receive your show-up fee.

Click "next" when you have finished reading to proceed to task 1.

Task 1

The first task is to solve some puzzles, a pattern game. On the screen, you will see a set of abstract pictures with one of the pictures missing. You need to choose a picture from the choices below to complete the pattern. For example, in the figure below, choice 8 completes the pattern and would be the correct answer.



(During the actual quiz, you will click on the correct answer, this is simply an example to show what it will look like)

You will have 10 minutes to complete this task. You will be asked to solve a total of 12 such puzzles. You will be paid for a random choice of two of these 12 patterns you will solve. For each correct choice, of the random two, you will receive £1. You can navigate through the

puzzles using the Next and Back buttons. Once you have completed all 12 puzzles, click the Finish button to submit your answers and complete the task.

Please ensure you do not click the Finish button before completing all 12 puzzles, as you will not be able to return to the task and may risk losing money.

If you have any questions, please raise your hand and we will come to help you. Please remain silent while we are running the exercise, as otherwise we will be forced to terminate the session!

Click "next" when you have finished reading to continue Task 1.

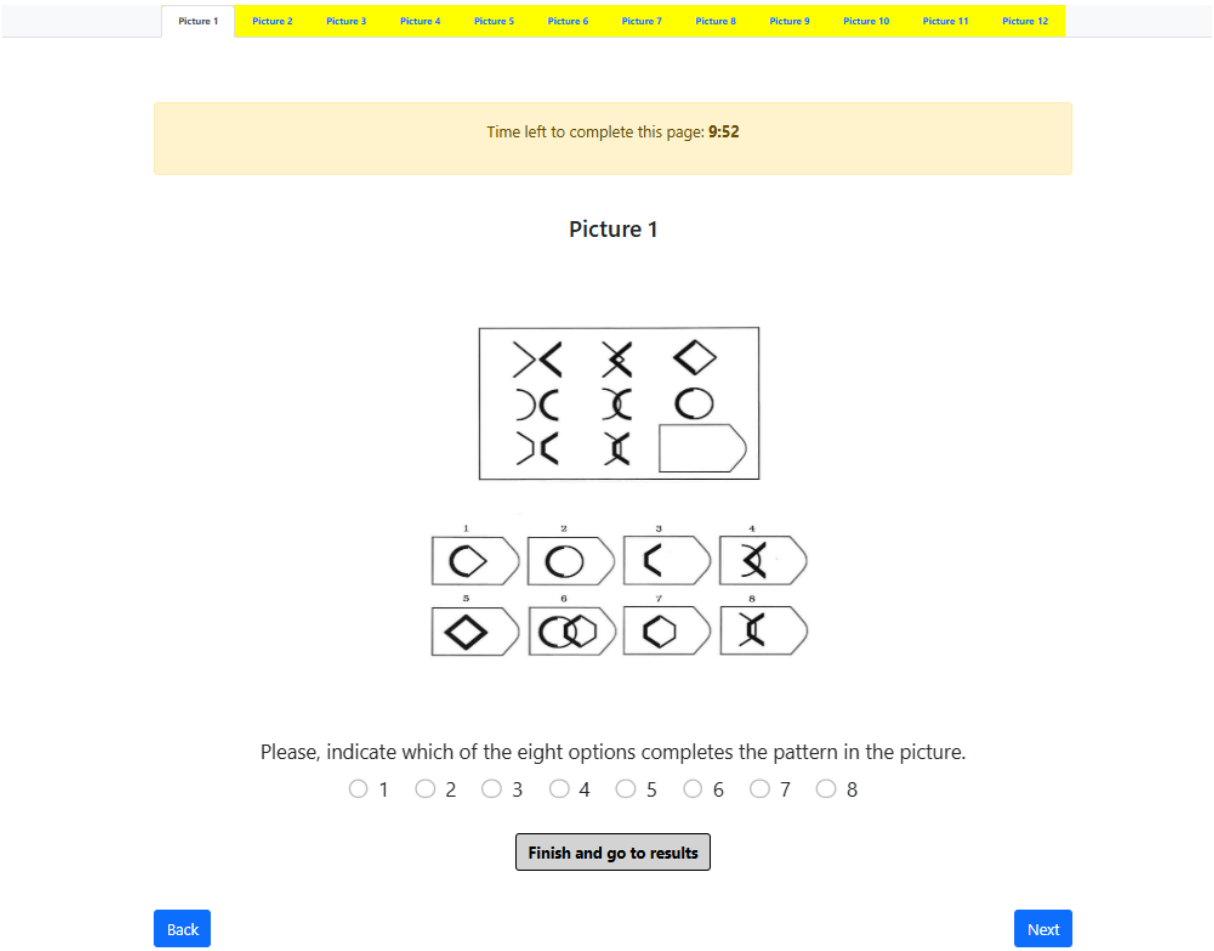


Figure A4: Screenshot of the implementation of Task 1

Task 2

Screen 1: This task is in the field of individual decision making. In this task you can earn real money. The amount you earn depends partly on your decisions and partly on chance. Please

read the instructions carefully, because a considerable amount of money is involved. You will have 10 minutes in total to read these instructions and complete two practice rounds. After that, the main task will begin. In the main task, you will have 15 minutes in total to complete all 25 rounds. During the experiment, we will refer to points instead of Pounds. Your earnings will be calculated in points and later paid to you in money.

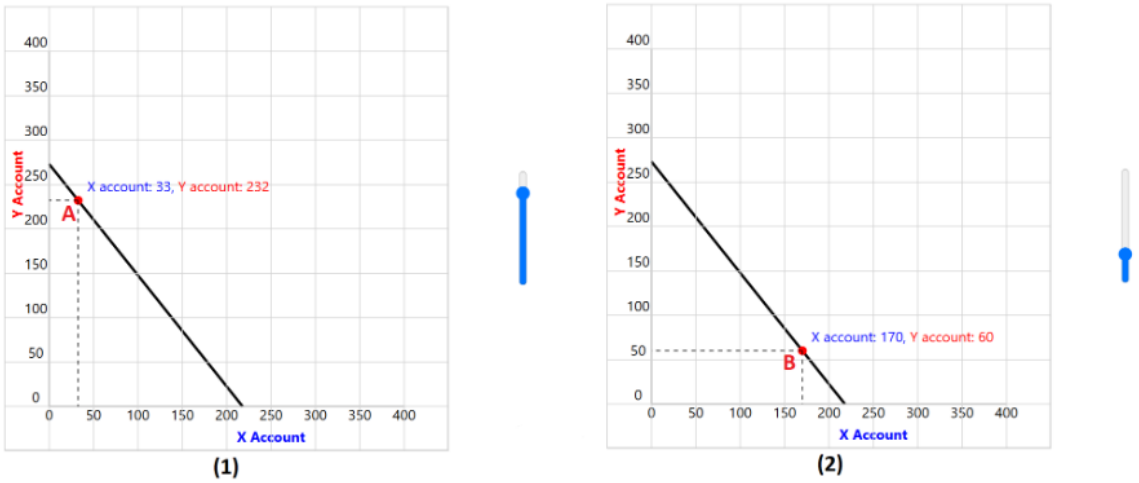
32 points = 1 Pound

This experiment consists of 25 rounds. Every round is independent, but your task will be the same in every round. In each round, your task is to distribute points between two accounts, called X account (Blue) and Y account (Red). At the start of each round, you will receive a set of possible distributions of points, out of which you are allowed to choose one. After you have made your decision, the computer will randomly select one of the accounts; both of the accounts X and Y have the same probability of being selected. Your earnings in a particular round are the points that you allocated to the account that has been selected by the computer, the points that you distributed to the other account do not count.

Screen 2: To make your decision in each of the rounds, you will use a chart. Below you can see what this looks like. The X account (Blue) corresponds to the horizontal axis, and the Y account (Red) corresponds to the vertical axis. Every point on the chart resembles a distribution between the X and Y account. At the start of each decision round, you will see a line which shows the possible distributions of points between the X and Y account. In each decision round you can only choose a point, which lies on the line. Examples of lines that you will see are shown below. In each round the computer selects one line which crosses both axes between 40 and 400 points, and at least one of the axes at 200 or more points. The lines that will be selected for you in the other decision rounds are independent of each other. In every round the computer will provide you with a new line. With every choice, you are allowed to choose any allocation of points between the X and Y account that is located on the line. You are allowed to choose one point.

An example of one of those lines is shown below. A possible choice is A, in figure (1), where

you allocate 33 points to the X account (blue) and 232 points to the Y account (red). Another possible choice is B, in figure (2), where you allocate 170 points to the X account (red) and 60 points to the Y account (b). There are many more points on this line that you can choose, other than point A and B alone. You are allowed to choose any point on the line. The only condition is that you can only choose one point in each round.



Screen 3: In the next screen you can practise two rounds. Click the practice button to begin. Then, you'll see a line with the possible distributions in this practice round. To choose a distribution, use the mouse to move the cursor on the screen along the line towards the distribution of your choice and then click on the left button on your mouse, to select your choice. (To do this, first move the cursor in the neighborhood of the point on the line.) You can see that you can only choose distributions that are located on the line. You can also use the slider located to the right of the figure to adjust the point along the line. Simply drag the slider to smoothly move the selection point to your preferred position. To activate the slider each time the page loads, first click on any point on the line.

You can click on any point on this line to see the allocated points to each account, blue and red. When you know which decision you would like to make, to confirm your decision, click the submit button. This button will only be visible once you click on a point. You will then be automatically transferred to the next round. After you have practised two rounds, click 'continue' to proceed to the next information page. This is a practice screen. The decision you make, will not be recorded.

Screen 4: Two practice rounds

Screen 5: After confirming your decision, the computer will randomly select one account; each of the two accounts, X and Y, have the same probability of being selected. If the X account is selected, then your earnings in this round equal the amount of points on the X account. The points on the other account are not used. If the Y account is selected, then your earnings in this round equal the amount of points on the Y account. The points on the other account are not used. Next, a new decision round starts. You will receive a graph with a new line from the computer. You will then make a new decision about the number of points to allocate to the X and the Y account. The computer will again choose between the X and Y account. This process repeats itself, until 25 rounds have passed. After the last round, you will see a screen showing a message that the experiment has ended.

At the end of the experiment (after round number 25), the computer randomly chooses one round from the 25 rounds. To do that, the computer draws a number between 1 and 25. So the selected round only depends on chance: all rounds have the same probability of being chosen. Because only one round (out of the 25 rounds in total) is randomly selected and paid out, there is no feedback after each round about the result (which round has been selected by the computer and the earnings in that round). The selected round, your own choice in that round and the amount that will be paid, are shown on the screen. The points will be exchanged to money. 32 points are worth 1 Pound.

The task starts now. From now, you will see 25 screens with the chart and 25 different lines. The computer chooses the lines based on chance. As explained earlier in the instructions, your task is to select a point from the possible point distributions. You can click on any point along the line. When you click on a specific point, you will be able to view the allocated points for each account associated with that selected point. To begin, you'll have to click the next button on the screen.

Screen 6: Task 2 Comprehension Questions

Please answer the following questions to check your understanding of the task instructions.

1. How is your payoff for a round determined?

    - By the total number of points you allocated across both accounts
    - By the number of points in the account randomly chosen by the computer
    - By the average of your allocations across X and Y

2. What is the probability that the computer chooses the X account in each round?

    - Higher than Y, it depends on how many points you allocate
    - Lower than Y, it depends on your choice
    - The same as Y, each account has a 50% chance

3. At the end of the experiment, how many of the 25 rounds are used to calculate your final payment?

    - All 25 rounds
    - Only 1 round, chosen at random
    - The last 5 rounds

Task 3

Screen 1 You will now participate in a different task. You will be matched randomly with another player. You and the other player are playing a game in which each player requests an amount of money. The amount must be (an integer) between 11 and 20 Experimental Pounds (EP). Each player will receive the amount he or she requests. A player will receive an additional amount of 20 Experimental Pounds if he or she asks for exactly one Experimental Pound less than the other player. 10 experimental pounds are equivalent to £1.

You will have 5 minutes in total to complete this task.

If for example, you request 19 EP and the other player requests 20 EP then,

Your payoff will be: 19 + 20 = 39EP (=£3.9)

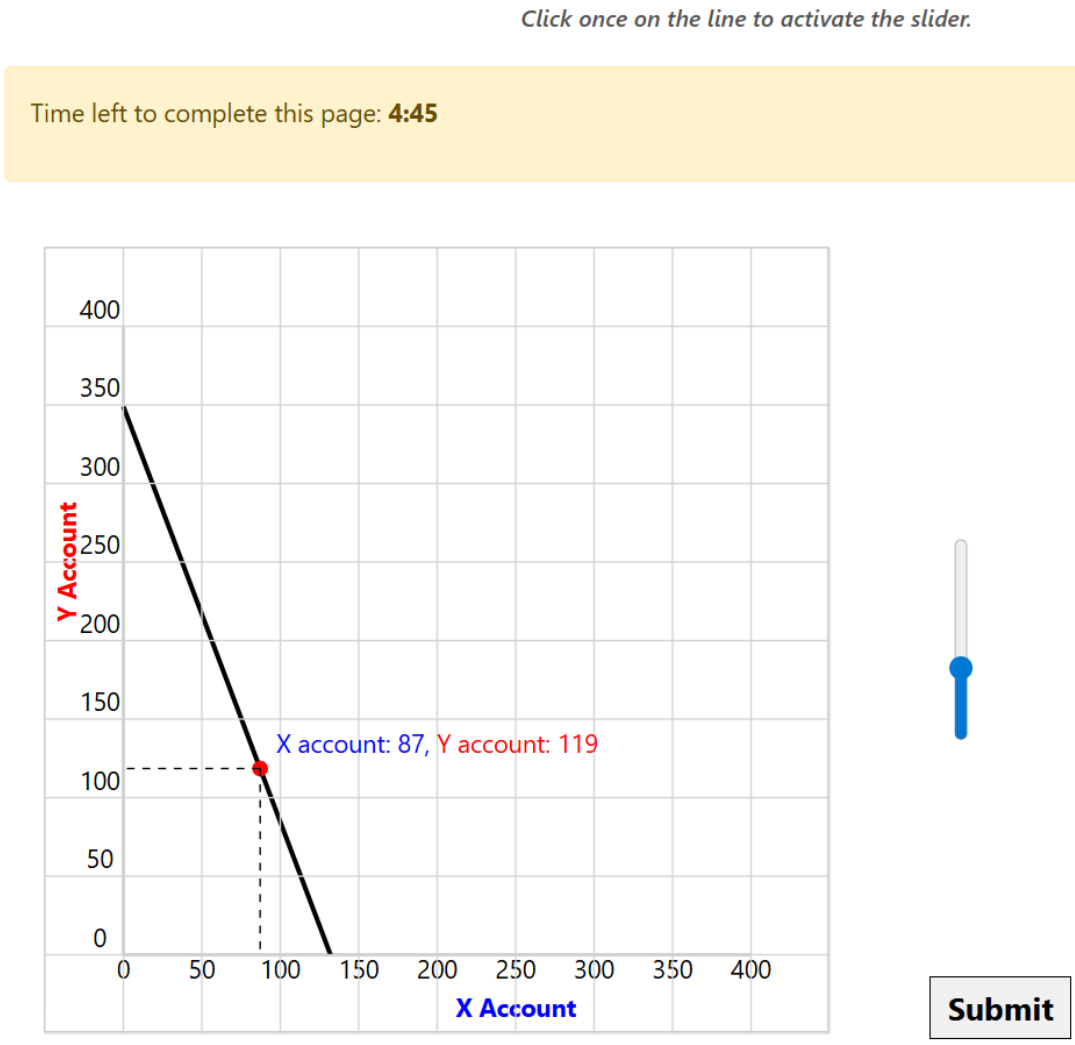


Figure A5: Screenshot of the implementation of Task 2

The other player's payoff will be: 20EP

If for example, you request 17 EP and the other player requests 16 EP then,

Your payoff will be: 17EP (=£1.7)

The other player's payoff will be: 16 + 20 = 36EP

If you have a question, please raise your hand.

If you have read the instructions and do not have any questions, please click 'Next' to proceed.

Screen 2: Make Your Request

Please enter an amount between 11 and 20 experimental pounds.

What amount of money would you request?

Screen 3: Please explain why you chose the amount you did:

Screen 4: Results

You requested: 14.

Your opponent requested: 14.

Your payoff: £1.40.

Your opponent's payoff: £1.40.

Task 4

In this task, each of you will be randomly matched with someone in this room to make decisions in several rounds.

On your screen you will see a similar screen like the one shown below. The computer will ask you to make a choice between C and D. Your payoff will be presented on the table, in green, and your partner's payoff will be presented on the right table, in black. In each table, your decisions (C or D) are represented in the rows and your partner's decisions are represented in the columns.

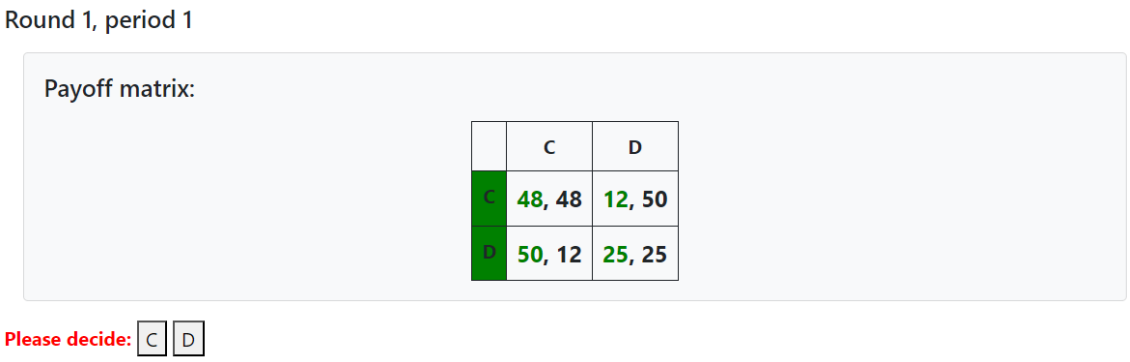


Figure A6: Example payoff tables for choices C and D.

The payoffs of each round will depend on both your decisions as well as your partner's. For

example, according to the table:

- If you choose C and your partner chooses C, your payoff will be 48 and your partner's payoff will be 48.
- If you choose D and your partner chooses C, your payoff will be 50 and your partner's payoff will be 12.
- If you choose C and your partner chooses D, your payoff will be 12 and your partner's payoff will be 50.
- And finally, if you choose D and your partner chooses D, your payoff will be 25 and your partner's payoff will be 25.

For each sequence of rounds (match) you will be randomly matched with someone from this room. This is done completely anonymously and no-one will ever know who you have been matched with.

After each round, there is a 75% probability that the match will continue for at least another round. That is, if there were 100 trials, in 75 of these the match would be repeated and in 25 the match would stop. So for example, if you are at the second round of the match, the probability there will a third round is 75% and similarly if you are at round 9, there will be a 75% probability for a further round. Once each match is finished, you will again be randomly matched with someone from this room and play the same sequence of rounds once again accordingly to the 75–25 probability. Whenever this happens you will be informed that a new match is starting.

The sum of the units that you will collect through all the matches, will determine your monetary payoff. Each unit corresponds to *0.4 pence*. Keep in mind that the game will be repeated many times and so you can potentially earn a lot of money!

Any questions? If you have any questions during the experiment, please raise your hand and we will come to help you. Please remain silent throughout the session as otherwise we will be forced to terminate the exercise and you will only receive your show-up fee.

Again, let us remind you that the length of each match is randomly determined. After each round, there is a 75% probability that the match will continue for at least another round. You will play with the same person for the entire match. Also, once a match is finished you will be randomly matched with another person for a new match.

Task 5

Screen 1: Task 5 is a choice task. On your screen you will see a list of 10 lottery choices. For each case you will be asked to indicate which of the two lotteries you would prefer to play. One out of these 10 lottery choices will be randomly picked and then the choice you've made will be played, and you will be paid according to the probabilities indicated.

If you have read the instructions and do not have any questions, please click 'Next' to proceed.

Screen 2:

Make a choice for each of the below rows. A random row will be chosen and you will be paid according to your choice for that row.

Option A	**Or...**	**Option B**
10% of £2.00, 90% of £1.50		10% of £3.50, 90% of £0.10
20% of £2.00, 80% of £1.50		20% of £3.50, 80% of £0.10
30% of £2.00, 70% of £1.50		30% of £3.50, 70% of £0.10
40% of £2.00, 60% of £1.50		40% of £3.50, 60% of £0.10
50% of £2.00, 50% of £1.50		50% of £3.50, 50% of £0.10
60% of £2.00, 40% of £1.50		60% of £3.50, 40% of £0.10
70% of £2.00, 30% of £1.50		70% of £3.50, 30% of £0.10
80% of £2.00, 20% of £1.50		80% of £3.50, 20% of £0.10
90% of £2.00, 10% of £1.50		90% of £3.50, 10% of £0.10
100% of £2.00, 0% of £1.50		100% of £3.50, 0% of £0.10

Task 6

Screen 1: Task 6 involves making intertemporal choices. On your screen, you will see a list of 14 choices. For each case, you'll be presented with two payment options: one for an early payment on the left and the other for a later payment on the right. One out of the 14 choices will be randomly picked to be played out.

If you have read the instructions and do not have any questions, please click 'Next' to proceed.

Screen 2:

Make a choice for each of the below rows. I would like to:

Option A	**Or...**	**Option B**
Receive £98 today		Receive £100 in 6 months.
Receive £94 today		Receive £100 in 6 months.
Receive £90 today		Receive £100 in 6 months.
Receive £86 today		Receive £100 in 6 months.
Receive £80 today		Receive £100 in 6 months.
Receive £70 today		Receive £100 in 6 months.
Receive £55 today		Receive £100 in 6 months.
Receive £98 in 6 months		Receive £100 in 12 months.
Receive £94 in 6 months		Receive £100 in 12 months.
Receive £90 in 6 months		Receive £100 in 12 months.
Receive £86 in 6 months		Receive £100 in 12 months.
Receive £80 in 6 months		Receive £100 in 12 months.
Receive £70 in 6 months		Receive £100 in 12 months.
Receive £55 in 6 months		Receive £100 in 12 months.

Task 7

Screen 2: Task 7 is a questionnaire. It is relevant to your background and personality. Your payment will not be affected by this. Again we would like to remind you that everything is anonymous so please answer as truthfully as possible as this is critically important for our research. If you have any questions, please raise your hand and we will come to help you.

If you have read the instructions and do not have any questions, please click 'Next' to proceed.

Screen 2:

Please select an answer indicating the extent to which you agree or disagree with that statement: "I see myself as someone who...".

	Disagree strongly	**Disagree a little**	**Neither agree nor disagree**	**Agree a little**	**Agree strongly**
Is talkative.					
Tends to find fault with others.					
Does a thorough job.					
Is depressed, blue.					
Is original, comes up with new ideas.					
Is reserved.					
Is helpful and unselfish with others.					
Can be somewhat careless.					
Is relaxed, handles stress well.					
Is curious about many different things.					
Is full of energy.					
Starts quarrels with others.					
Is a reliable worker.					
Can be tense.					
Is ingenious, a deep thinker.					
Generates a lot of enthusiasm.					
Has a forgiving nature.					
Tends to be disorganized.					
Worries a lot.					
Has an active imagination.					
This is not a real personality question but an attention check please answer “Agree strongly” so we know you are paying attention.					

(continued)

	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
Tends to be quiet.					
Is generally trusting.					
Tends to be lazy.					
Is emotionally stable, not easily upset.					
Is inventive.					
Has an assertive personality.					
Can be cold and aloof.					
Perseveres until the task is finished.					
Can be moody.					
Values artistic, aesthetic experiences.					
Is sometimes shy, inhibited.					
Is considerate and kind to almost everyone.					
Does things efficiently.					
Remains calm in tense situations.					
Prefers work that is routine.					
Is outgoing, sociable.					
Is sometimes rude to others.					
Makes plans and follows through with them.					
Gets nervous easily.					
Likes to reflect, play with ideas.					
Has few artistic interests.					
Likes to cooperate with others.					
Is easily distracted.					
Is sophisticated in art, music, or literature.					

Task 8

Screen 1: Task 8 which is the last section for today is a very quick questionnaire.

Please click ‘Next’ to proceed.

Screen 2:

Demographic Questions

What is your age?

Year of study? 1 2 3 4 other

What is your gender? Female Male Other Prefer not to say

What is your country of origin?

What is your mother tongue?

What was the highest possible mark in your high school?

What was your final mark?

What is your current degree course?

Would you consider your course mostly: Quantitative or Qualitative?

Quantitative

Qualitative

Have you solved a similar pattern game (like the one today) in the past year?

Yes

No

F.1 Experiment Implementation Details

Table A8: **Maximal period (T) of each SG.**

SG	T
1	1
2	4
3	2
4	2
5	1
6	2
7	12
8	4
9	4
10	5
11	8
12	2
13	1
14	7
15	2
16	4
17	4
18	1
19	4
20	1
21	5
22	7
23	3
24	1
25	1
26	4
27	1
28	1
29	2
30	2

References

Afriat, Sidney N (1972), "Efficiency estimation of production functions." *International economic review*, 568–598.

Alaoui, Larbi, Katharina A. Janezic, and Antonio Penta (2020), "Reasoning about others' reasoning." *Journal of Economic Theory*, 189, 105091.

Alaoui, Larbi and Antonio Penta (2016), "Endogenous depth of reasoning." *The Review of Economic Studies*, 83, 1297–1333.

Alós-Ferrer, Carlos and Johannes Buckenmaier (2021), "Cognitive sophistication and deliberation times." *Experimental Economics*, 24, 558–592.

Arad, Ayala and Ariel Rubinstein (2012), "The 11-20 money request game: A level-k reasoning study." *American Economic Review*, 102, 3561–73.

Basteck, Christian and Marco Mantovani (2018), "Cognitive ability and games of school choice." *Games and economic behavior*, 109, 156–183.

Carvalho, Leandro S, Stephan Meier, and Stephanie W Wang (2016), "Poverty and economic decision-making: Evidence from changes in financial resources at payday." *American economic review*, 106, 260–284.

Charness, Gary, Aldo Rustichini, and Jeroen Van de Ven (2018), "Self-confidence and strategic behavior." *Experimental Economics*, 21, 72–98.

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

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv (2007), "Consistency and heterogeneity of individual behavior under uncertainty." *American economic review*, 97, 1921–1938.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman (2014), "Who is (more) rational?" *American Economic Review*, 104, 1518–50.

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

Eber, Nicolas, Patrick Roger, and Tristan Roger (2024), "Finance and intelligence: An overview of the literature." *Journal of Economic Surveys*, 38, 503–554.

Gill, David and Victoria Prowse (2016), "Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis." *Journal of Political Economy*, 124, 1619–1676.

Goeree, Jacob K, Philippos Louis, and Jingjing Zhang (2018), "Noisy introspection in the 11–20 game." *The Economic Journal*, 128, 1509–1530.

Guan, Menglong and Ryan Oprea (2026), "Three faces of complexity in strategic choice."

Halevy, Yoram, Dotan Persitz, and Lanny Zrill (2018), "Parametric recoverability of preferences." *Journal of Political Economy*, 126, 1558–1593.

Holt, Charles A and Susan K Laury (2002), "Risk aversion and incentive effects." *American economic review*, 92, 1644–1655.

John and Jean Raven (2003), "Raven progressive matrices." In *Handbook of Nonverbal Assessment* (R Steve McCallum, ed.), 223–237, Springer US, Boston, MA.

John, Oliver P, Eileen M Donahue, and Robert L Kentle (1991), "Big five inventory." *Journal of personality and social psychology*.

Lambrecht, Marco, Eugenio Proto, Aldo Rustichini, and Andis Sofianos (2024), "Intelligence disclosure and cooperation in repeated interactions." *American Economic Journal: Microeconomics*, 16, 199–231.

Moghaddasi Kelishomi, Ali and Daniel Sgroi (2024), "A field study of donor behaviour in the iranian kidney market." *European Economic Review*, 170, 104887.

Nielsen, Kirby and John Rehbeck (2022), "When choices are mistakes." *American Economic Review*, 112, 2237–2268.

Polisson, Matthew, John K-H Quah, and Ludovic Renou (2020), "Revealed preferences over risk and uncertainty." *American Economic Review*, 110, 1782–1820.

Proto, Eugenio, Aldo Rustichini, and Andis Sofianos (2019), "Intelligence, personality, and gains from cooperation in repeated interactions." *Journal of Political Economy*, 127, 1351–1390.

Proto, Eugenio, Aldo Rustichini, and Andis Sofianos (2022), "Intelligence, errors and cooperation in repeated interactions." *Review of Economic Studies*, 89, 2723–2767.

Sofianos, Andis (2025), "Intelligence in experimental economics." *Oxford Research Encyclopedia of Economics and Finance*.

Stango, Victor and Jonathan Zinman (2023), "We are all behavioural, more, or less: A taxonomy of consumer decision-making." *The Review of Economic Studies*, 90, 1470–1498.

Stanovich, Keith E (2012), "On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning." In *The Oxford Handbook of Thinking and Reasoning* (Keith J Holyoak and Robert G Morrison, eds.), Oxford University Press, Oxford, United Kingdom.