Three Faces of Complexity in Strategic Choice*

Menglong Guan† Ryan Oprea‡

February 28, 2026

Abstract

We experimentally show that complexity shapes strategy selection in infinitely repeated prisoner's dilemmas in several fundamental ways. Our subjects choose from a menu of pre-set strategies and we vary the presence of three potential drivers of complexity aversion by varying whether subjects have to (i) implement their chosen strategies themselves, (ii) compute the payoff consequences of those strategies and whether they (iii) can use the simplicity of strategies as a coordination device. We find that subjects display a strong aversion to algorithmically complex strategies when these drivers of complexity aversion are present, but not when they are removed. In our data, complexity is at least as important in shaping strategic choice as strategic uncertainty, the most important determinant of strategy selection identified by the literature so far. We show that these results are robust to learning and several measures of cognitive sophistication.

Keywords: complexity, strategic uncertainty, repeated games, repeated prisoner's dilemmas

*We are grateful to Guillaume Fréchette, Alistair Wilson, Emanuel Vespa and Sevgi Yuksel for valuable conversations and comments that informed this research. All mistakes are our own. This research was approved by the Institutional Review Board of Penn State. Menglong Guan acknowledges financial support from the Smeal College of Business, Penn State. Ryan Oprea acknowledges financial support from the National Science Foundation under Grant SES-1949366.

†Guan: Smeal College of Business, The Pennsylvania State University, mkg6150@psu.edu.

‡Oprea: Haas School of Business, University of California, Berkeley, roprea@gmail.com.

1 Introduction

The folk theorem (Friedman 1971, Fudenberg & Maskin 1986, Fudenberg, Levine & Maskin 1994) describes one of the fundamental indeterminacies in economic theory: with sufficiently patient players, many distinct strategies can be supported in equilibrium in infinitely repeated games. As a result, repeated game theory often fails to make sharp predictions of how people will behave in even simple settings like the repeated prisoner's dilemma. This indeterminacy has given rise to a number of efforts to discover principles of equilibrium selection that can be used to make repeated game theory more predictive.

One of the most important theoretical efforts in this direction is the "automata literature," which observes that (i) repeated game strategies differ in their *complexity* and (ii) that distaste for complex strategies (due to, e.g., cognitive effort costs) might shrink the set of strategies supportable in equilibrium (Aumann 1981, Neyman 1985, Rubinstein 1986, Abreu & Rubinstein 1988, Sabourian 2004).[1] To model this, the automata literature describes repeated game strategies as algorithms (i.e., finite automata), performed not by computers but by human agents. The literature shows that if players attach even arbitrarily small costs to the complexity of these algorithms (e.g., the number of states or transitions in their algorithmic structure), the equilibrium set will shrink, yielding sharper predictions. This work therefore hypothesizes that aversion to *implementing* complex strategies is an important driver of strategy choice in repeated games and that accounting for these implementation costs might make game theory more predictive.

The aim of our paper is to empirically study the role of this "implementation complexity" in strategy selection in repeated games, and to compare that influence to the influence of other potential drivers of complexity aversion and other principles of equilibrium selection.

On the one hand, as the *theoretical literature* has emphasized, complex strategies are not only difficult to implement: they are also difficult to *compute*. As strategies become more complex (as their algorithmic structure expands), the cost of understanding their payoff implications increases under standard worst case analysis from computer science (Gilboa 1988, Ben-porath 1990). This suggests that complex strategies may not only be distasteful to implement (as implementation complexity sug-

[1] See Kalai (1990) and Chatterjee & Sabourian (2009) for surveys of the literature.

gests), but also may produce uncertainty about the payoff implications of strategies due to 'computational complexity." Aversion to such uncertainty produces a second reason (in addition to implementation costs) that players might prefer to avoid complex strategies.

On the other hand, as the *experimental literature* has emphasized, strategy selection in repeated games is a coordination game, characterized by significant strategic uncertainty.[2] Arguably, the most important finding from the extant experimental literature on repeated games is that strategic risk (as measured by the size of the basin of attraction of strategies) is a powerful predictor of strategy selection. In particular, subjects tend to avoid cooperation increasingly as cooperative strategies become more strategically risky (Dal Bó & Fréchette 2011, 2018, Blonski, Ockenfels & Spagnolo 2011, Blonski & Spagnolo 2015, Calford & Oprea 2017, Embrey, Fréchette & Yuksel 2017, Vespa & Wilson 2017, Dal Bó, Fréchette & Kim 2021, Boczoń, Vespa, Weidman & Wilson 2024). The complexity of strategies may interact with this driver of strategy choice, for the simple reason that a great deal of literature in psychology shows that simplicity is *salient* and therefore plausibly a tool for coordination in strategic settings.[3,4] As a result, in addition to influencing strategy selection via implementation costs and computational failures, complexity may serve as a coordinating device for resolving strategic uncertainty because simple strategies are *salient*. This is a third way complexity may influence strategy choice.

In order to measure and compare the role these three facets of complexity (implementation, computation and salience) play in strategy choice, we report an experiment (using nearly 600 subjects and 9000 strategy choices) designed to study how subjects choose between strategies in infinitely repeated prisoner's dilemmas, the workhorse environment for studying repeated game strategies in experimental economics. In traditional experiments in this literature, subjects make decisions naturalistically by making action choices in each round of indefinite implementations of the prisoner's dilemma. However, it is difficult to pose sharp tests of the influence of

[2]In particular, as the folk theorem shows, infinitely repeated games (with sufficiently patient players) generates a large number of equilibria involving a great many possible strategies. Because of this, in the space of strategies, infinitely repeated games are effectively coordination games.

[3]The search for simplicity is considered a fundamental principle in human cognition (Chater 1999, Chater & Vitányi 2003) and perception (Feldman 2016).

[4]Simplicity was considered a determinant of focal points by Schelling: "the intrinsic magnetism of particular outcomes, especially those that enjoy prominence, uniqueness, *simplicity*, precedent, or some rationale that makes them qualitatively differentiable from the continuum of possible alternatives" (Schelling 1960, 70).

complexity in such settings, given the extremely large strategy space in these games. Instead, we follow a smaller recent strand of the literature (Dal Bó & Fréchette 2019, Cason & Mui 2019, Gill & Rosokha 2024) by having subjects directly choose strategies from carefully controlled menus that allow for tight tests of these claims.[5]

Specifically, in each of the 15 consecutive prisoner's dilemma supergames of our experiment, we ask pairs of randomly matched subjects to choose from a menu of seven strategies (including always defect, grim trigger, tit-for-tat, and their more complex variations), and use the strategy pairs they choose to determine their payoffs. In a baseline treatment (called *All*), subjects have to both (i) infer the payoff consequences of strategy pairs (meaning they face computational complexity) and (ii) implement their chosen strategy themselves after selecting it (meaning they face implementation complexity). In the *ComputeSalience* treatment (computational complexity but no implementation complexity), the computer automatically implements the strategy the subject selects, removing implementation complexity from the problem while retaining computational complexity.[6] In the *ImplementSalience* treatment (implementation complexity but no computational complexity), subjects are told the payoff consequences of all strategy combinations prior to choice, retaining implementation complexity but removing computational complexity from the problem. By comparing how subjects choose strategies across these three treatments, we can measure the impacts of implementation and computational complexity on strategy choice. Finally, in the *Salience* treatment subjects face neither implementation nor computational complexity but still have descriptions of strategies available to coordinate on, while in the *None* treatment, these descriptions (and their potential coordinating effects) are removed. By comparing *Salience* to *None*, we can further measure the role the salience of simple strategies plays in strategy selection.

We find that complexity has a strong influence on strategy choice, shaping it in fundamental ways. Subjects' likelihoods of choosing algorithmically simple strategies (strategies with fewer states and transitions) fall by more than one-half (and in some comparisons as much as four-fifths) when all three drivers of complexity aversion are removed, and standard metrics of the algorithmic complexity of chosen strategies

[5]There are also experimental studies, including Embrey, Mengel & Peeters (2017), Dal Bó & Fréchette (2019) and Romero & Rosokha (2018, 2019, 2023), that allow subjects to construct their own strategies.

[6]Specifically, subjects are incentivized to implement their strategy against a computer-implementation of their counterpart's strategy (with errors in the computer's implementation that are based on estimates from subjects' error rates in prior data).

nearly double. What's more, we find that each of these three drivers of complexity aversion individually has a significant impact on strategy choice, driving subjects towards the use of algorithmically simpler strategies. These complexity effects are strong, leading subjects to sacrifice 44% of their potential earnings. In order to better understand the magnitude of these effects, we compare the impact of complexity on strategic choice to the impact of strategic uncertainty, the most important driver of strategy choice identified by the literature so far. By design, our experiment includes strategy comparisons in which standard responses to strategic uncertainty predict the opposite choice as aversion to complex strategies. In these contests between the two drivers of strategic choice, we find that computational and implementation complexity dominate the influence of strategic uncertainty, suggesting that complexity may be at least as important as strategic uncertainty in shaping strategy choice.

Our design also allows us to assess the robustness of these results in several ways, giving us a clearer understanding of the underlying mechanism. Most importantly, we find that the effects of complexity are robust to experience and feedback—indeed, if anything, we find that the impact of complexity strengthens over the course of the 15 supergames subjects play. Similarly, we find that the effects of complexity are largely robust to measures of cognitive ability, with some effects actually strengthening for more cognitively able subjects. Likewise, we find that results are virtually unchanged when conditioning on high performance in comprehension questions, suggesting that these results are robust to the attentiveness of subjects. Finally, we find similar results in lab and online sessions, suggesting that these results are largely robust to the use of a plausibly more sophisticated subject pool.

These findings harmonize with and deepen findings from the prior experimental literature on repeated games. One of the striking findings from this literature is that subjects tend to select very simple strategies, a regularity that is not well-explained by previous empirical findings. In standard repeated prisoner's dilemmas, for instance, three ultra-simple strategies (always defect, grim trigger, and tit-for-tat) account for over 70% of the observed strategies in most studies (Dal Bó & Fréchette 2018). While strategic uncertainty is important in this prior literature for understanding cooperation, it is important primarily in governing the choices subjects make between an already small and simple set of strategies. Our findings help to explain why this menu is so small, and also suggest that subjects are unlikely to deviate very far from this menu even if it were valuable for mitigating strategic uncertainty.

These findings also connect to a growing empirical literature on the role complexity plays in economic behavior (Oprea 2026). Most closely related is Oprea (2020), which directly measures the costs of implementing algorithms describable as finite automata, and shows that automata states and transitions (the complexity metrics we use in our analysis) generate significant implementation costs. Also related is Banovetz & Oprea (2023), which shows that removing the cognitive difficulties associated with automata state complexity, alters the procedures subjects use in non-strategic bandit games. Jones (2014) show that subjects make less cooperative choices in repeated games in which cooperation requires the use of more complex strategies – a finding that is highly consistent with ours.[7] In a concurrent study, He (2022) studies questions similar to ours using a very different design, and finds some evidence that implementation complexity influences strategy choice while arguing that computational complexity is more important in his setting. Further afield are recent papers showing that complexity has first order influences on behavior in lottery choice (Bernheim & Sprenger 2020, Puri 2025, Enke & Graeber 2023, Oprea 2024, Enke & Shubatt 2023), intertemporal choice (Enke, Graeber & Oprea 2025), information demand (Guan, Oprea & Yuksel 2025), formation of beliefs (Ba, Bohren & Imas 2025), responses to taxes (Abeler & Jäger 2015), and mechanism design (Jakobsen 2020).

Finally, our findings inform the automata literature, which theoretically characterizes the effects complexity has on strategic behavior. This literature typically (and conservatively) models complexity costs as infinitesimally small and shows that even these tiny costs can have significant effects on behavior and on the equilibrium set. By comparison, our evidence suggests that the costs of implementing algorithms are far from infinitesimal and are instead quite significant relative to game payoffs. What's more, computational complexity and salience have a powerful additional influence that intensify aversions to algorithmically complex strategies. The significant costs implied by our findings suggest that these models can be made substantially more restrictive and therefore more predictive by assuming more substantial aversion to complexity, justified by evidence like ours. Indeed, we believe justifying the assumption of much larger costs in models of complexity in games is among our paper's most

[7] Wilson & Wu (2017) studies how adding the option of terminating the repeated game (arguably an extremely simple strategy) affects cooperation and efficiency in the repeated prisoner's dilemma. They find that subjects overuse termination and get inefficient outcomes for a range of termination payoffs – a result that could be explained by subjects economizing on the implementation costs of continuing the repeated game and selecting actions round by round.

important contributions.

The remainder of the paper is organized as follows. Section 2 introduces the conceptual background. Section 3 describes the experimental design. Section 4 presents the main results. Section 5 concludes and discusses the implications of our results.

2 Conceptual Background

Our study focuses on the infinitely repeated prisoner's dilemma. In each of an infinite number of periods, two players simultaneously choose actions A (defection) or B (cooperation) from the game matrix visualized in Table 1 and earn the payoffs visualized in the cells of the matrix from the combination of choices. Payoffs are discounted in each period of the game by a discount factor δ.

Table 1: The Prisoner's Dilemma

	A	B
A	0, 0	8, -2
B	-2, 8	6, 6

Notes: Actions A and B correspond to "defect" and "cooperate", respectively. The same game is used by Fudenberg, Rand & Dreber (2012) and Cason & Mui (2019). The payoffs are in experiment tokens. Each token is worth $0.5 in our experiment.

Formally, let $g = (\Gamma, u)$ be a symmetric two-player game in normal form (e.g., a standard prisoner's dilemma) where Γ is the action set and $u : \Gamma^2 \to \mathbb{R}$ is the payoff function. Let $G = (g, \delta)$ be the infinitely repeated game with the stage game g and discount factor δ. In G, two players play the stage game g over an infinite horizon with payoffs discounted by δ. Let $\gamma_t^i \in \Gamma$ be player i's action choice in round t and H be the history of actions with $h_t = \{(\gamma_1^i, \gamma_1^j), ..., (\gamma_{t-1}^i, \gamma_{t-1}^j)\}$. A strategy $\theta : H \to \Gamma$ specifies the action choice in each round t given h_t. Let Θ be the set of available strategies from which players can choose. Anticipating our experimental design, the game can then be described as $\Omega = (G, \Theta)$, in which two players choose simultaneously a strategy $\theta \in \Theta$ that determines what actions they will take in each round of the repeated game G to respond to actions in past rounds, and the payoff is the one induced in G by the pair of chosen strategies. Let $\pi_i(\theta_i, \theta_j)$ be the (expected) repeated game payoff to player i. The pair of strategy (θ_i, θ_j) is a *Nash Equilibrium* in the game Ω if there is no θ_i' or θ_j' such that $\pi_i(\theta_i', \theta_j) > \pi_i(\theta_i, \theta_j)$ or $\pi_j(\theta_i, \theta_j') > \pi_j(\theta_i, \theta_j)$.

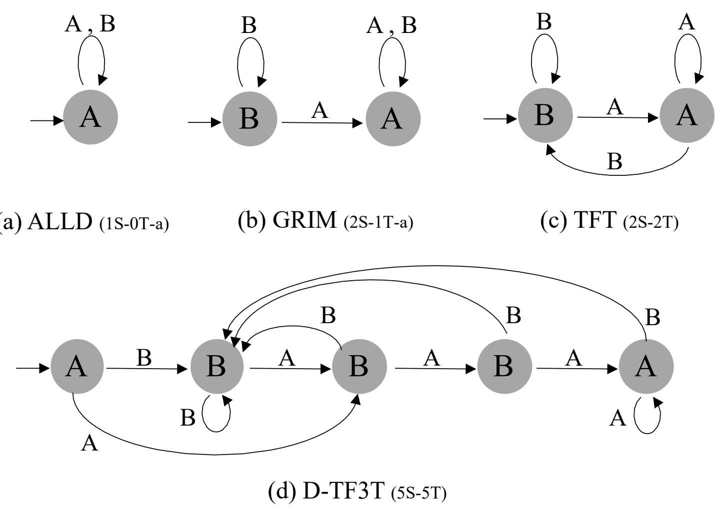

Figure 1: Automaton representations of Strategies *Notes: Notations in the parentheses denote the complexity of strategies in terms of their automaton representations. xS-yT denotes that the number of states (s-complexity) of a strategy is x and the number of transitions in excess of the number of states (t-complexity) is y; a indicates the strategy has absorbing states.*

2.1 Strategies as Algorithms and Measures of Complexity

Our goal is to understand the role *complexity* plays in people's choice of strategies in games like the repeated prisoner's dilemma just described. To study this, we follow the "automata literature" (Aumann 1981, Neyman 1985, Rubinstein 1986, Abreu & Rubinstein 1988) which (following the typical definition in computer science) defines "complexity" as the cost of implementing an algorithm (i.e., a rule). The literature formalizes this idea by describing repeated game strategies as finite automata (finite-state machines), simple formal descriptions of algorithms imported from computer science.[8] A finite automaton is a four-tuple (S, s^0, f, τ). S is a set of states, s^0 is the initial state, $f : S \to R$ is an output function that specifies a response in each state, and $\tau : S \times E \to S$ is a transition function that specifies a next state as a function of the current state and event. Figure 1 shows the automaton representations of three commonly observed strategies in the repeated prisoner's dilemma, ALLD (always defect), GRIM (grim trigger), and TFT (tit-for-tat), and a more complex strategy D-TF3T (exploitative tit-for-three-tat).[9] The gray circles are states, the letters inside the circles are actions taken in the corresponding state (so $R = \{A, B\}$), the arcs are transitions between states, the letters next to arcs are events (actions by counterparts

[8] Kalai & Stanford (1988) show that any finite strategy in an infinitely repeated game can be described by a finite automaton.

[9] Table 6 in Appendix A provides detailed descriptions of all strategies studied in our experiment.

in the previous period) triggering those states (so $E = \{A, B\}$), and the left-hand arrow indicates the initial state s^0. For example, the GRIM automaton instructs players to begin playing action B (in the left-most state), to continue with this action as long as the counterpart has played B in the prior period, but to transit to the right-hand state if the counterpart ever plays A, and then (as the right-hand state instructs) to play A forever.[10]

The automata literature hypothesizes that metrics of the *size* of an automaton measure the complexity of the strategy it describes (i.e., the cognitive costs associated with the strategy). The size of an automaton is usually measured in the literature by the number of *states*, $|S|$, in the strategy's automaton representation (sometimes called s-complexity in the literature) though it is sometimes instead measured by the number of *transitions* in excess of the number of states (sometimes called t-complexity in the literature). Thus, ALLD, GRIM, TFT and D-TF3T in Figure 1 have s-complexities of 1, 2, 2 and 5 respectively and t-complexities of 0, 1, 2 and 5. Oprea (2020) directly elicits the costs of implementing finite state machines in an individual choice setting and provides evidence that both states and transitions generate complexity costs, as the literature has suggested. He also finds that the total number of transitions (the number of states plus the number of net transitions) parsimoniously summarizes the complexity costs of finite state machines and therefore serves as a good overall complexity measure. We will therefore generally order the complexity of strategies by the total number of transitions in their automaton representations (which is equivalent to s-complexity plus t-complexity). Under this usage, GRIM (3 total transitions) is more complex than ALLD (1 total transition), TFT (4 total transitions) is more complex than either of the two, while D-TF3T (10 total transitions) is the most complex of all.

The automata literature hypothesizes that the cognitive costs of *implementing a strategy* increase in the complexity of the automaton describing that strategy. That is, strategies with more states and transitions are more subjectively costly to implement and therefore, on the margin, less likely to be selected.[11] Following the literature, we

[10] To be clear, the strategies described throughout the paper are always defined over implemented (realized) actions, i.e., the actions actually taken in past rounds, regardless of whether they resulted from intentional choices or implementation errors. This is the standard definition in both the theoretical automata literature and the experimental literature on repeated games (including experiments with implementation errors).

[11] Importantly, "implementation complexity" as we (and the automata literature) use the term refers to the subjective cognitive costs of the mental effort required to implement a strategy, *not* to

will call this first potential driver of complexity aversion, "implementation complexity".

Notion 1 (Implementation Complexity): *Implementing complex strategies is subjectively costly (due, e.g., to the costs of the mental effort required), reducing players' willingness to use them.*

The automata literature has used this notion to improve the predictiveness of repeated game theory. For instance, Rubinstein (1986) and Abreu & Rubinstein (1988) study a game in which each player chooses (and commits to) a strategy (described by a finite automaton) in two-person infinitely repeated games and the player's preference depends both on the average repeated game payoffs and the cost of implementation complexity (as measured by s-complexity) of the player's strategy. They show that introducing even a very small complexity cost (very small costs attached to automaton states) dramatically reduces the set of equilibrium payoffs.[12,13,14] Banks & Sundaram (1990) introduces the measure of t-complexity and proposes a criterion for comparing the complexity of strategies, which combines s-complexity and t-complexity. The paper shows that the set of equilibria in the setting of Abreu & Rubinstein (1988) further shrinks under the proposed criterion for assessing complexity.

The automata literature has also highlighted a second sense in which complexity might influence strategy choice. Complex strategies are not only difficult to implement, they are also potentially difficult to reason about. In particular, as the strategy of another player grows increasingly complex, it becomes increasingly difficult to identify how to best respond to that strategy because the computations involved become more burdensome. Formally, as Gilboa (1988) and Ben-porath (1990) show, the problem of identifying what strategy to play as a best response to another's strategy is *NP complete*, meaning the computational burden of the problem explodes as the strategies

the payoff losses players fear may stem from mistakes in properly implementing their strategies.

[12]Specifically, they focus mainly on the case where implementation costs enter preferences lexicographically: a player first identifies the best-response strategies; if multiple best responses exist, the player chooses the one with the lowest s-complexity.

[13]Neyman (1985) points out that a similar notion can justify cooperation in finitely repeated prisoner's dilemmas. He does this by restricting the available strategies of players to those with s-complexity that does not exceed a certain upper bound and shows that a cooperative equilibrium can be supported in the finitely repeated prisoner's dilemma for a fixed upper bound of s-complexity and a sufficiently large number of repetitions.

[14]Following Rubinstein (1986) and Abreu & Rubinstein (1988), Chatterjee & Sabourian (1998, 2000), Sabourian (2004), and Gale & Sabourian (2005, 2006) study equilibrium refinement based on complexity costs of implementing strategies/automata in dynamic matching and bargaining games.

under consideration become more complex in the sense discussed above.[15,16] These computational problems only compound when players have to evaluate the payoff consequences of their own strategy choice, which requires reasoning about how their strategy interacts with each possible opponent strategy.[17]

We will call this second source of complexity, following the literature, "computational complexity":[18] it is difficult for subjects to assess the payoff implications of their strategy choices, making them *uncertain* about these implications. There are at least two behavioral implications of this kind of complexity-derived uncertainty. First, we should expect this kind of uncertainty to reduce players' likelihoods of best responding to the strategy choices of others. Second (and more importantly for our purposes), as recent research suggests, this kind of complexity-derived uncertainty can make more complex strategies less attractive to players.[19] Computational complexity (like implementation complexity) therefore may produce a systematic aversion to complex strategies.

Notion 2 (Computational Complexity): *Evaluating the payoff consequences of*

[15] In computer science and computational complexity theory, NP complete stands for "nondeterministic polynomial-time complete" and is a classification of a set of problems. An NP complete problem is considered one of the hardest problems in NP, meaning finding a polynomial-time solution for any NP complete problem would imply a polynomial-time solution for all problems in NP.

[16] Gilboa (1988) shows that computing the best response automaton of a player given several other players' automata in repeated games is NP-complete. The problem becomes polynomial when the number of players is fixed. Ben-porath (1990) shows that computing a best response automaton in a two-person repeated game when there is uncertainty about the opponent's automaton is NP-complete. The problem becomes polynomial if the size of the opponent's strategy space is known in advance. Papadimitriou (1992) shows designing an automaton with bounded s-complexity to best respond to a given automaton in the repeated prisoner's dilemma is NP-complete. The problem becomes polynomial (in the s-complexity of the given automaton) if no bound is imposed on the s-complexity of the response strategy.

[17] Note that the original theoretical literature frames computational complexity as the difficulty of computing a best response to a given (pure or mixed) strategy of the opponent. In our setting, the relevant problem is somewhat bigger: subjects, when choosing their strategy, must reason about how their strategy interacts with each possible opponent strategy. Clearly, the two problems are closely related—evaluating the payoff consequences of available strategies against all possible opponent strategies generalizes the problem of computing a best response to any particular one—and both become harder as the algorithmic complexity of the strategies under consideration increases.

[18] Computational complexity can be interpreted as another layer of implementation complexity: it can be thought of as the implementation complexity of an evaluation procedure needed to compute the best response. More discussions can be found in Murawski & Bossaerts (2017) and Oprea (2020).

[19] de Clippel, Moscariello, Ortoleva & Rozen (2025) provides experimental evidence that complexity produces cognitive uncertainty (Enke & Graeber 2023) about the value of menu options and that ambiguity averse subjects respond to this uncertainty by *avoiding* complex options. Thus, complexity-derived uncertainty produces systematic *complexity aversion*. We hypothesize that computational complexity may have a similar effect in strategy choice.

complex strategies is difficult, producing subjective uncertainty that reduces players' willingness to use complex strategies.

2.2 Strategic Uncertainty and Complexity

The experimental literature on repeated games has gathered significant evidence that *strategic uncertainty* plays a central role in guiding and influencing strategy choice. The infinitely repeated prisoner's dilemma is a coordination game in the space of strategies, with multiple equilibria, saddling players with significant strategic risk whenever they elect to use cooperative strategies. The literature (Dal Bó & Fréchette 2011, 2018, Blonski et al. 2011, Blonski & Spagnolo 2015, Calford & Oprea 2017, Embrey, Fréchette & Yuksel 2017, Vespa & Wilson 2017, Dal Bó et al. 2021, Boczoń et al. 2024) has operationalized this idea by calculating the *basin of attraction* (the BOA) of strategies (often over pairs, e.g., GRIM versus ALLD) and using its size as a metric of the degree of strategic risk a player exposes herself to when choosing cooperative strategies. Formally, for a 2×2 strategy choice game as in Table 2, the BOA has a well-defined closed-form expression: $BOA_s = \frac{\alpha-\eta}{\alpha-\eta+\beta-\rho}$ (and $BOA_c = \frac{\beta-\rho}{\alpha-\eta+\beta-\rho}$). A strategy is risk-dominant (Harsanyi & Selten 1988) and therefore conventionally considered low in strategic risk if its BOA is larger than one-half. Using this measure, the literature has gathered robust evidence that a strategy is more likely to be chosen if the size of its BOA is larger (i.e., more risk dominant or more robust to strategic risk).

Table 2: Example of Strategy Choice Game

	strategy s	strategy c
strategy s	$\alpha,\ \alpha$	$\rho,\ \eta$
strategy c	$\eta,\ \rho$	$\beta,\ \beta$

Notes: Let $\alpha > \eta$ and $\beta > \rho$, then (s, s) and (c, c) are two pure-strategy equilibria of the game.

Our experiment considers strategy choice games with larger strategy spaces—beyond 2×2—where closed-form BOA expressions are not directly applicable and the interaction structure among multiple strategies is considerably richer (Golman & Page 2010). Following Sandholm (2010) and Proto, Rustichini & Sofianos (2022), we compute the BOA of each available strategy in a large game by simulating the evolution of strategy frequencies in a population under best response dynamics. The best response dynamic is formally defined as:

$$\forall s \in S : \frac{d\mu(t,s)}{dt} = \lambda(\tau(s) - \mu(t,s)), \text{ for some } \tau \in BR(\mu(t))$$

where $\mu(t,\cdot) \in \Delta(S)$ represents the population state (frequency distribution over pure strategies $s \in S$), λ is a fixed parameter controlling the speed of adjustment, and the best response correspondence is:

$$BR(\mu) = \{\tau \in \Delta(S) : \pi(\tau,\mu) \geq \pi(s,\mu) \text{ for all } s \in S\}$$

where $\pi(s,\mu)$ denotes the expected payoff from playing strategy s against population state μ. Operationally, we first generate a large set of random initial population states $\mu(0)$ uniformly distributed on the simplex $\Delta(S)$; for each initial state, simulate the best response dynamics until convergence; classify the limiting state as converging to strategy s if $\mu(\infty, s)$ approaches 1; and finally compute the BOA of strategy s as the proportion of initial states that converge to s.[20]

Strategic uncertainty is important for our purposes for two reasons. First, it suggests a third and final role complexity might play in governing strategy choice. Because strategy choice is a coordination game beset by strategic uncertainty, it can be influenced by the salience of choice options as described by Schelling (1960). That is, if strategies have prominent or salient features (sometimes referred to as "labels" in the literature), players may use these features as coordination devices in choices of strategies. A literature in psychology (Chater 1999, Chater & Vitányi 2003, Feldman 2016) suggests that simplicity itself is often a salient characteristic, raising the possibility that simple strategies serve as coordinating devices simply because of their simplicity. This is a third, distinct way that complexity may influence strategy choice:[21]

Notion 3 (Salience): *Simple strategies are more salient than complex strategies, allowing them to serve as coordinating devices in the face of strategic uncertainty.*

A second reason strategic uncertainty is important for understanding complexity, is that its prominence in the prior literature makes it a natural benchmark for as-

[20]As a validation check, we find this simulation-based approach can successfully recover the closed-form BOA values of strategies in 2×2 games.

[21]Our notion of salience can also be interpreted through the lens of the idea of prominence-driven salience described by Bordalo, Gennaioli & Shleifer (2022): simple strategies attract bottom-up attention because they are more accessible to human cognition and perception and thus are more salient.

sessing the magnitudes of complexity effects. Directly comparing the effects of the types of complexity sketched above to the BOA, particularly when they make conflicting predictions about strategy preferences, will allow us to assess the relevance and strength of complexity effects relative to the most important forces identified in the literature so far.

2.3 Discussion

It is important to emphasize that these three putative drivers of complexity aversion are hypotheses and that it is far from obvious, ex ante, that they (individually or collectively) will have the effect of driving players towards the use of significantly simpler strategies. For one thing, it is not clear, ex ante, that all of these notions will generate complexity aversion. For instance, computational complexity may lead to mere noisiness rather than systematic aversion and it is possible that simplicity is not the most important principle of salience used in coordinating strategy. For another, it is not obvious, ex ante, that these potential drivers will be strong enough to appreciably influence choice, especially relative to competing drivers of strategic choice like aversion to strategic risk or payoff differences across strategy combinations. Finally, there is the potential that complexity could have other effects that encourage rather than discourage the selection of complex strategies, overriding these effects. For instance, Robson (2003) describes settings in which evolutionary forces drive players to use more complex rather than simpler strategies and it is possible that related forces influence strategy choice in settings like ours. For these reasons, direct empirical examination of these hypotheses is valuable and for this purpose we designed the experiment described in the next section.

3 Experimental Design

Our experiment implements a variation on the infinitely repeated prisoner's dilemma described in Section 2.[22] Relative to the discussion there, we make three changes that aid us in answering our motivating questions:

- **Indefinite Repetition:** Because implementing an infinitely repeated game is impractical in an experiment, we follow the bulk of the literature by implement-

[22] An earlier version of this paper reported a somewhat different experiment focused on imperfect monitoring games. We re-designed and re-ran the experiment in response to a set of excellent comments by referees. Findings from that earlier experiment are very similar to those from the current experiment. See Online Appendix D for details.

ing a theoretically isomorphic *indefinitely repeated game* (Fréchette & Yuksel 2017). Instead of being a discount factor, δ is a per-period probability that the game continues into another period.

- **Explicit Strategy Selection:** Because it is important for our questions to carefully control the strategies available to subjects, we (i) follow a recent literature (Dal Bó & Fréchette 2019, Cason & Mui 2019, Gill & Rosokha 2024) by asking subjects to choose from a curated menu of available strategies[23] and (ii) pay subjects the expected value of the resulting strategy profile (described to subjects as the average outcome over one million implementations of the strategy profile).[24]

- **Ecological Trembles:** An artificiality of explicit strategy selection designs like ours is that they impose perfect *implementation* of the strategies subjects select. By contrast, estimates of subjects' strategy use in naturalistic prisoner's dilemma experiments suggest that they tremble (make errors) each round with a small likelihood (on the order of 5 or 6% at the median). To make explicit strategy selection more realistic, we therefore impose a small rate of error, estimated from actual subject behavior,[25] when calculating the payoffs of strategy profiles. We do not tell subjects this error rate, but only that trembles will occur with a "small chance (based on estimates from previous participants behavior in games in which they directly chose their actions each round)," forcing subjects to (as they must in naturalistic prisoner's dilemmas) form their own beliefs based on their beliefs about real subjects' noisiness.

In the experiment, a subject chooses their strategy from a menu, knowing that a human counterpart is selecting an (unknown) strategy from the same menu. Because the subject does not know what strategy their counterpart will select, they face *strategic uncertainty* when selecting a strategy.

[23]These experiments tend to find broadly similar results as more conventional designs and several strands of evidence suggest they accurately measure strategy preferences (Dal Bó & Fréchette 2019).

[24]In naturalistic designs in which subjects originate their own strategies "on the fly" while playing multiple periods of the game, controlling the set of strategies available is impossible given the infinite size of the strategy space.

[25]In particular, we used the strategy implementation error rate estimated by Fudenberg et al. (2012) who, in one of their treatments, study the same stage game and continuation probability as we do in a naturalistic prisoner's dilemma experiment (i.e., in which subjects choose actions round-by-round rather than explicitly selecting strategies). They estimate error rates of 5% per round in their subjects, leading us to include implementation errors of 5% per round when simulating the payoffs of strategy profiles.

Finally, the experiment varies the presence of the three potential drivers of complexity aversion discussed in Section 2, which we manipulate in the design:

- **Implementation**: In some variations of the problem, after selecting a strategy, subjects are asked to implement the strategy against their counterpart's strategy over an indefinite horizon, requiring them to bear any subjective costs implementation produces. In others, they are not required to implement their strategy themselves, removing implementation complexity from the task.
- **Computation**: In some variations of the problem, subjects have to infer or compute the payoffs stemming from combinations of strategies prior to selecting a strategy. In others, we explicitly show subjects the payoffs stemming from each combination, removing computational complexity from the task.
- **Salience**: In some variations of the problem, subjects are given a description of the strategy they (and their counterpart) are choosing from the menu of available strategies, allowing the simplicity of the strategy to serve as a coordinating device for strategy selection. In others, strategy descriptions are replaced with meaningless identifiers in the strategy menu, thus removing the coordinating potential of simplicity from the task.

Table 3: Summary of Treatments

Treatment \ **Variation**	salience	computation	implementation
All	✓	✓	✓
ComputeSalience	✓	✓	
ImplementSalience	✓		✓
Salience	✓		
None			

Notes: A check mark means the scope for the corresponding effects of complexity exists in the treatment.

These treatment variables are arrayed in the design in a way summarized in Table 3. The *All* treatment is designed to be closest to the standard infinitely repeated prisoner's dilemma game, where subjects have to infer or compute the payoff consequences of strategies and have to implement their strategies round by round themselves. The comparison between the *All* and *ComputeSalience* (*ImplementSalience*) treatments allows us to measure whether and to what extent implementation (computational) complexity affects strategy choices as does the comparison between *Salience* and *ImplementSalience* (*ComputeSalience*). The former comparison measures the effect of one driver of complexity aversion in the presence of the other, while the latter captures the effect in the absence of the other. More generally, the four treatments that include salience—*All*, *ComputeSalience*, *ImplementSalience*, and *Salience*—form a 2×2 factorial design crossing computational and implementation complexity, allowing us to study not only the main effects of each driver but also their interaction.[26] Finally, the contrast between *Salience* and *None* treatments isolates the effect of the salience of simple strategies on strategy selection.

3.1 The Strategy Choice Game

Table 4: The Strategy Choice Game

	ALLD 0 (1S-0T-a)	GRIM 0 (2S-1T-a)	TFT 0 (2S-2T)	D-GRIM2 0.0024 (4S-3T-a)	D-TF2T 0.2485 (4S-4T)	D-GRIM3 0.1094 (5S-4T-a)	D-TF3T 0.6397 (5S-5T)
ALLD	2	10	12	9	13	15	20
GRIM	1	28	30	8	16	14	22
TFT	0	27	37	8	39	35	40
D-GRIM2	1	9	17	35	37	36	38
D-TF2T	0	7	38	35	40	38	40
D-GRIM3	-1	8	36	35	39	40	40
D-TF3T	-2	6	39	35	40	40	40

Notes: Numbers in cells are the payoffs of the row player given corresponding strategy combinations. The payoffs of the column player are symmetric. The number below each strategy is the size of the basin of attraction (BOA) of the strategy. Notations in the parentheses denote the complexity of strategies in terms of their automaton representations. xS-yT denotes that the number of states (s-complexity) of a strategy is x and the number of transitions in excess of the number of states (t-complexity) is y; a indicates the strategy has absorbing states. The payoffs of strategy choices are computed based on 1,000,000 simulations given the stage game as shown in Table 1, $\delta = 7/8$, and a small (5%) chance that each player might deviate from her strategy each round, and rounded to whole numbers. The payoffs are in experiment tokens. Each token is worth \$0.5 in our experiment.

In each of 15 rounds of the experiment, subjects face a strategy choice game that uses (i) the stage game, g, shown in Table 1, and (ii) continuation rate $\delta = \frac{7}{8}$. In each round, subjects are given a menu of 7 strategies, Θ={ALLD, GRIM, TFT, D-GRIM2, D-TF2T, D-GRIM3, D-TF3T}. Table 4 presents the payoffs of each strategy combination based on one million simulations given the stage game g and continuation rate δ. We designed this strategy choice game with several considerations in mind. First, ALLD, GRIM and TFT are by far the most commonly observed strategies in standard repeated prisoner's dilemma experiments, making up 70% of strategies observed in most experimental data (Dal Bó & Fréchette 2019). Second, we

[26]Clearly, we cannot similarly assess interactions between salience and the other two potential drivers of complexity (implementation and computational complexity): in order to impose implementation or computational complexity, we must *necessarily* describe strategies to subjects, immediately producing scope for salience effects.

expanded this list with more complex versions of GRIM and TFT to create variations in the complexity level of strategies and their BOAs. Third, given that complex strategies have relatively higher BOAs (and the most complex strategy has the highest BOA) and are therefore less strategically risky, this game features an inherent tension between complexity and strategic uncertainty in many pairwise comparisons, allowing us to examine which effect dominates when they make opposite predictions. Finally, the payoffs from coordinating on the same strategy (numbers on the diagonal of Table 4) are relatively flat, which helps to control the impact of other confounding factors such as payoff dominance, allowing us to more cleanly measure the influence of complexity.[27,28]

3.2 Treatment Variations

We studied this strategy selection game under a series of between-subjects treatments (i.e., all subjects are assigned to only one treatment) designed to vary the availability of each of the three potential drivers of complexity aversion (*implementation*, *computation* and *salience*) discussed in Section 2.

Varying Computational Complexity. In all treatments, after the subject chooses her strategy, she receives in payment a *base payment* equal to the expected value of her chosen strategy given her counterpart's chosen strategy. Figure 2 shows a screenshot of the strategy menu from the *All* and *ComputeSalience* treatments in which subjects are shown their counterpart's menu of rules in the column header and asked to select from the same menu of strategies shown as rows.[29] However, in these treatments, subjects are not told the payoff consequences of these strategy combinations (represented by question marks in the table): in both treatments, subjects must compute

[27]Indeed, our decision to feature exploitative strategies in our strategy menu (i.e., D-GRIM2, D-TF2T, D-GRIM3 and D-TF3T rather than GRIM2, TF2T, GRIM3 and TF3T) was entirely because they produce flatter payoffs across symmetric strategy pairs and therefore better control payoff dominance. One potential concern may be that featuring exploitative strategies introduces an additional distinction between non-exploitative and exploitative strategies that might confound inferences about complexity aversion. Appendix C.3 shows that the treatment effects of all three facets of complexity (discussed below) remain robust when we restrict the analysis to comparisons among exploitative strategies only.

[28]There remain some payoff differences for symmetric strategies, meaning that payoff dominance may still play a role. However, if so, we should expect this to work *against* our hypotheses regarding the influence of complexity because coordinating on complex strategies tends to lead to slightly higher payoffs in the game. Indeed, we will use these payoff differences in Section 4.3 to coarsely assess how much money subjects are willing to sacrifice to avoid complex strategies.

[29]For each subject, the order of the seven available strategies is randomized and then fixed across the fifteen matches they experience.

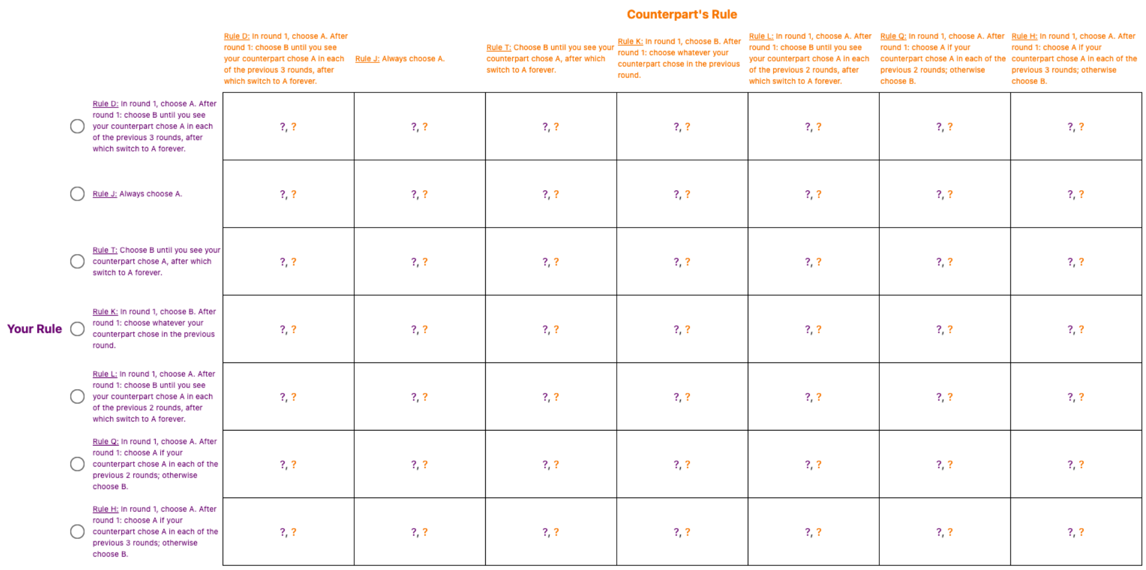

Figure 2: Interface of Strategy Choice in All and ComputeSalience Treatments *Notes: Question marks mean that subjects must compute or infer payoff consequences of these strategy combinations themselves. In treatments without computational complexity (i.e., ImplementSalience, Salience, and None), the question marks are replaced by actual payoff matrices as shown in Table 4. In treatments without implementation complexity, there is a reminder message saying "The computer will automatically perform your chosen rule for you against your counterpart's rule." below the submit button on this page. For each subject, the order of the seven available strategies is randomized and then fixed across fifteen matches. In the experiment, subjects can mouse over any rule to see it in a larger font. See more details in Appendix E.*

or infer the payoff consequences themselves, exposing them to the full computational complexity of the task. By contrast, in the *ImplementSalience*, *Salience* and *None* treatments, subjects are explicitly shown the relevant payoff numbers from Table 4 in the cells of the same matrix, removing computational complexity from the game.

Varying Implementation Complexity. Although subjects in all treatments are given a base payment equal to the expected value of their chosen strategies in all treatments, in the *All* and *ImplementSalience* treatments, subjects nonetheless are required to *implement* their strategy against their counterpart's strategy (implemented by the computer) in a single realization of the game. Specifically, in these treatments we included an Implementation Stage that follows each of their 15 strategy choices (i.e., in each of the 15 consecutive games). Figure 3 shows a screenshot from these treatments. After selecting her strategy, the subject is shown, one by one, action realizations of her counterpart's chosen strategy and is asked to implement the exact strategy she selected in response. Doing this forces the subject to experience the implementation costs of her chosen strategy. By contrast, in treatments *ComputeSalience*, *Salience* and *None* the subject does not have to implement her chosen strategy, removing the complexity of implementation from the game.

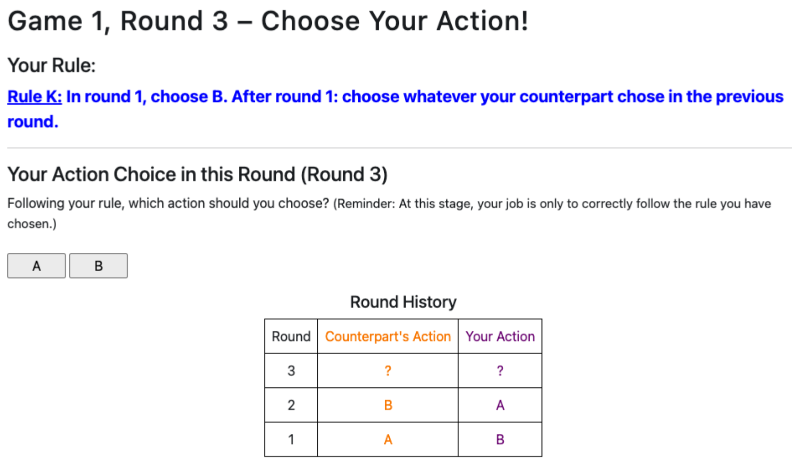

Figure 3: Interface of the Implementation Stage in All and ImplementSalience Treatments

Subjects' choices in the Implementation Stage do not impact the base payments from their strategy choices which (as in all treatments) are determined by the strategy choice alone. However, in order to motivate subjects to actually implement her chosen strategy (and thus bear the relevant mental effort costs of implementation) we impose a small penalty for each mistake the subject makes in implementing exactly the strategy she has selected.[30] Thus, in the *All* and *ImplementSalience* treatments, the subject is incentivized to internalize the subjective effort costs of the strategy she chose. Importantly, because the penalty is small and (ex post) the rate of implementation errors is modest (under 4% per round) these penalties have very little impact on overall payoffs (less than 2% of maximum payoffs from strategy choices, even for the most algorithmically complex strategies)[31] and therefore are highly unlikely to have a significant influence on strategy choices in these treatments.[32] Instead, treatment effects from requiring subjects to implement their strategies are primarily driven by the subjective effort costs of implementing complex strategies.

[30]Specifically, a subject loses one experiment token ($0.5) from her base payment every time her action choice does not follow what her chosen strategy specifies. Our data strongly suggests that the penalty scheme is sufficient to induce subjects to implement their chosen strategies: among all subjects from the treatments with implementation complexity, on average, the rate at which subjects made mistakes in implementing their chosen strategies each a round is only 3.7%.

[31]Ex post, we calculated: (1) the average rate at which subjects make implementation errors (per round) when implementing each strategy, (2) the expected number of mistakes (and thus expected penalty costs) for each strategy given the average game length of $1/(1-\delta)$. We find that expected penalty costs are economically negligible, with the maximum being 0.65 experiment tokens across all strategies which is very small relative to the maximum possible payoff of 40 tokens.

[32]In particular, the modest expected monetary losses from making mistakes in implementing strategies is unlikely to be a major driver of strategy choice in these treatments.

As discussed above, one artificiality of Strategy Choice designs like ours is that they artificially ensure that subjects' strategies are implemented without error. By contrast, in typical structural estimates of strategy choice (e.g., studies discussed in Table 10 of Dal Bó & Fréchette (2018)), estimated error rates in the implementation of strategies range from 2.7% to 11.5%, with a median between 5% and 6%. Given our motivating questions, this artificiality is consequential: much of the implementation complexity of strategies only arises off-path, meaning perfect implementation of strategies can be expected to artificially attenuate the implementation complexity subjects face (e.g., in our *All* and *ImplementSalience* conditions). To remove this artificiality, we informed subjects that in the Implementation Stage their counterpart's strategy will be implemented wrongly (i.e., with error) with a 'small chance...based on estimates of previous participants' behavior" each round. We deliberately did not tell the subject what this probability was (the number we used, based on previous estimates, was 5% each round)[33] so that (as in real games) subjects had to form their own expectations about the likelihood of counterpart error. For consistency, we include the same rate of error (for both players) in the simulations we use to calculate subjects' base payments (shown in Table 4) and inform subjects of this fact.[34]

Varying the Salience of Simple Strategies. Finally, in the *None* treatment, we remove the *description* of the strategies from the strategy choice game, meaning the relative complexity of strategies cannot even be used as a coordination device. Figure 4 shows a screenshot. Unlike other treatments, strategies in the None treatment are described using identifiers (letters) that give no sense of the relative complexity of the strategies available. This removes the potential for the simplicity of strategies to serve the coordinating role it may play in the other treatments. Thus, by comparing strategy choice in the *Salience* and *None* treatments, we can examine the role this third facet of complexity plays in strategic choice.

[33] This estimate is taken from Fudenberg et al. (2012) who study the same stage game and continuation probability as we do and estimate an error rate of 5% in structural estimates of strategies (see their Table A2). This number is also similar to the error rate of 3.7% that we (ex post) measured in our own Implementation Stages.

[34] Specifically in the instructions describing the calculation of the base payment we wrote "In order to make this more realistic...we will include a small chance (based on estimates from previous participants' behavior in games in which they directly chose their actions each round) that you and your counterpart accidentally choose the wrong action each round." When describing the Implementation Stage to subjects the instructions included the following language: "For your counterpart's actions, we will again include a small chance (based on estimates of previous participants' behavior) each round that your counterpart's rule might be wrongly implemented."

		Counterpart's Rule						
		Rule Q	Rule J	Rule D	Rule K	Rule H	Rule L	Rule T
	Rule Q	40, 40	0, 13	38, 39	38, 39	40, 40	35, 37	7, 16
	Rule J	13, 0	2, 2	15, -1	12, 0	20, -2	9, 1	10, 1
	Rule D	39, 38	-1, 15	40, 40	36, 35	40, 40	35, 36	8, 14
Your Rule	Rule K	39, 38	0, 12	35, 36	37, 37	40, 39	8, 17	27, 30
	Rule H	40, 40	-2, 20	40, 40	39, 40	40, 40	35, 38	6, 22
	Rule L	37, 35	1, 9	36, 35	17, 8	38, 35	35, 35	9, 8
	Rule T	16, 7	1, 10	14, 8	30, 27	22, 6	8, 9	28, 28

Figure 4: Interface of Strategy Choice in the None Treatment *Notes: Each rule corresponds to one of the seven available strategies, but subjects are not shown the descriptions. For each subject, the order of rules is randomized and then fixed across fifteen matches.*

3.3 Repetition and Matching

Subjects played the same game (presented in Table 4) 15 times and were randomly matched with a (possibly different) counterpart each time.[35] Subjects received feedback on their payoffs after each game, allowing for possible game-to-game learning as is common in repeated prisoner's dilemma experiments. Our primary data was collected using Prolific, a popular online subject pool. Because of difficulties with conducting experiments on multi-player simultaneous games using online subject pools like Prolific, we implemented the experiment as an individual decision-making experiment, but in a way that retains strategic uncertainty. To do this, we first ran a lab session of each treatment in which randomly paired participants played the strategy choice game simultaneously in the LEMA laboratory at Penn State.[36] Prolific subjects were then matched with the lab session participants (specifically, with those participants' chosen strategies). The matching in each treatment was pre-determined randomly and applied to all Prolific subjects in the same treatment.[37] As we show in

[35] To maximize the comparability of treatments, the game lengths (number of rounds) were predetermined randomly according to the geometric distribution with p=1-δ and applied to subjects in all treatments. Following Fudenberg et al. (2012), we generated a sequence of random game lengths according to the geometric distribution and then picked a starting point in that sequence.

[36] The lab sessions are identical to Prolific sessions except that participants in the lab played the strategy choice game simultaneously. The lab session of All, ComputeSalience, ImplementSalience, Salience, and None treatments had 18, 20, 18, 18, and 20 participants, respectively. Lab participants were randomly paired with each other in each game. Each lab session was scheduled for 75 minutes, and participants typically completed the experiment within 50 minutes.

[37] In the instructions, we informed Prolific subjects that their counterparts' strategy choices were collected earlier. Please refer to Appendix E for more details.

Section 4.5.4, results from the lab and Prolific experiments are broadly similar.

3.4 Incentives and Implementation Details

We ran all of the treatments on Prolific with 500 subjects, using software programmed by the authors in oTree (Chen et al. 2016). The experiment was pre-registered on AsPredicted.org. Each subject participated in only one treatment, and each treatment consisted of 100 total subjects.[38] Subjects were given instructions electronically and were given comprehension questions that they had to answer correctly before proceeding – over 90% of subjects answered these questions correctly on the first try and results are unchanged if we condition on only these subjects (see Section 4.5.3).[39] Subjects received \$7 for participation and an extra \$0.5 if they answered all comprehension questions correctly on their first attempt. 20% of the subjects in each treatment were randomly selected to receive an additional bonus payment based on their decisions in a randomly selected game and in the cognitive test questions. The average bonus payment (among the 20% subjects) is about \$18.6. The median time spent on the experiment was about 35, 26, 38, 28, and 23 minutes in *All*, *ComputeSalience*, *ImplementSalience*, *Salience*, and *None* treatments, respectively. At the end of the experiment, we also administered ten cognitive questions, including one Wason selection task, three cognitive reflection test questions, three Raven's test questions, and three belief bias syllogisms, which we discuss in more detail in Section 4.5.2.[40]

4 Results

4.1 Complexity and Strategic Choices

We begin by quantifying the overall effect complexity has on strategic choice by comparing our *All* treatment (in which all three facets of complexity are present to influence behavior) and our *None* treatment (in which none are). To do this, we study

[38] We recruited subjects who are based in the US, between 18 and 65 years of age, fluent in English, have a Prolific approval rate above 99%, and have not participated in the earlier experiments we conducted in 2022-2023. At the end of the experiment, we asked participants whether they had used any LLM tools to help them with the experiment. For each treatment, data collection continued until we reached 100 participants who reported no use of any LLM tools. The pre-registration can be found at https://aspredicted.org/xx5nc7.pdf.

[39] In all treatments, subjects answer two comprehension questions: one on how the game length is determined and one on how their base payment is determined. In treatments with implementation complexity (*All* and *ImplementSalience*), subjects answer a third question on how their total earnings from a game are determined. See details in Appendix E.

[40] These questions are the same as those in Oprea & Yuksel (2022) and Guan et al. (2025).

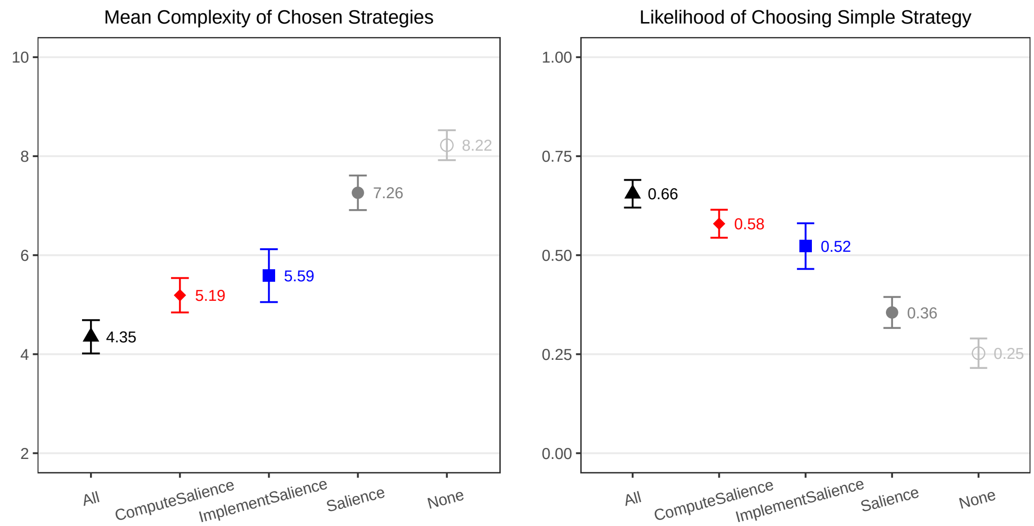

Figure 5: Complexity Level of Chosen Strategies and Likelihood of Choosing Simple Strategy *Notes: The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. When computing the likelihood of choosing (the relatively) simple strategy in the right panel, a subject's strategy choice in a game is analyzed as six pairwise comparisons. For both panels, we first compute the statistic for each subject and then calculate the treatment-level averages. The short vertical lines denote 95 percent confidence intervals.*

two statistics specified in our pre-registration, pooling over all fifteen supergames of the experiment.[41]

First, the left panel of Figure 5 plots the average algorithmic complexity level (i.e., states plus transitions in the automaton representation) of the strategies subjects choose in each treatment.[42] From treatment *All* to *None*, as each of the three faces of complexity is removed from the task, the average complexity level of chosen strategies nearly doubles from 4.35 to 8.22 (Wilcoxon rank sum test, $p < 0.001$).[43] Thus subjects become strongly averse to algorithmically complex strategies when the three drivers of complexity aversion discussed in Section 2 are present.

Second, the right panel of Figure 5 plots subjects' likelihood of choosing the relatively (algorithmically) simple strategy in pairwise comparisons.[44] From treatment

[41]In the pre-registration, we wrote "Our main analyses will focus on the last 5 games. Additionally, we will analyze all 15 games and compare behaviors in early versus late games to explore learning effect." We slightly deviate from this (purely for expositional purposes) by beginning with an analysis of the data overall and discussing learning in its own section. We provide analyses of only the last five matches in Appendix C.1. The main results remain the same.

[42]The complexity levels of strategies ALLD, GRIM, TFT, D-GRIM2, D-TF2T, D-GRIM3, and D-TF3T are 1, 3, 4, 7, 8, 9, and 10, respectively.

[43]We also analyze the median complexity level of chosen strategies and find similar results. Details are presented in Figure 12 of Appendix B.

[44]This is computed as follows: a subject's choice of a strategy over the other six available strategies is analyzed as six pairwise comparisons, each containing a relatively simple strategy and a relatively

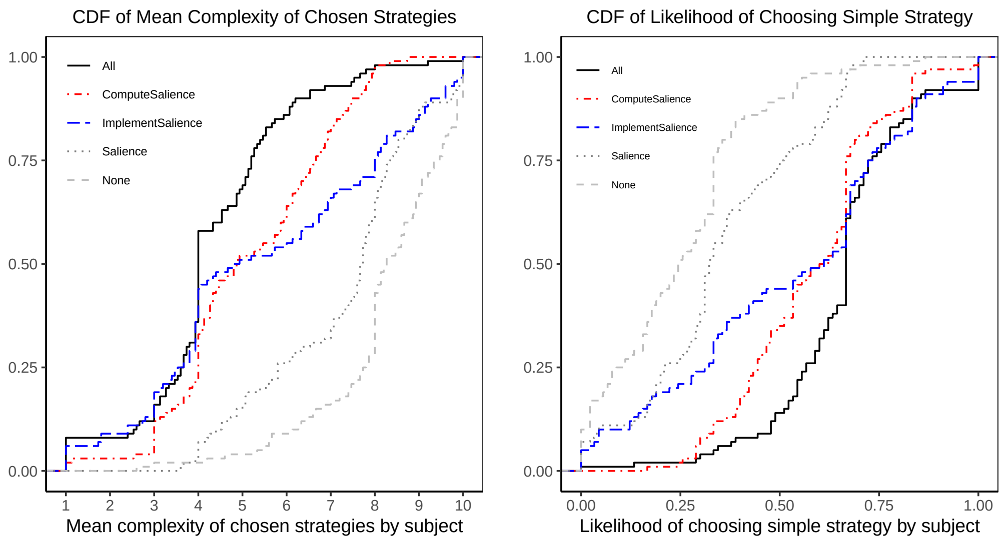

Figure 6: Distributions of Complexity Level of Chosen Strategies and Likelihood of Choosing Simple Strategy *Notes: The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. When computing the likelihood of choosing (the relatively) simple strategy in the right panel, a subject's strategy choice in a game is analyzed as six pairwise comparisons.*

All to *None*, the likelihood of choosing the simpler strategy falls sharply by more than half from 0.66 to 0.25 ($p < 0.001$). Indeed, as Figure 13 in Appendix B shows, from treatment *All* to *None*, the empirical distribution of chosen strategies displays a substantial shift from relatively simple strategies (i.e, ALLD, GRIM, and TFT) to the more complex ones. Again, this suggests that subjects are far more averse to algorithmically complex strategies when they are exposed to the three drivers of complexity aversion hypothesized in Section 2.

Importantly, these patterns are not driven by outliers but are broadly represented in the distribution of subjects. Figure 6 plots treatment-wise empirical CDFs of the average complexity of chosen strategies by subject (left panel) and subject-wise likelihoods of choosing the algorithmically simpler strategy in pairwise comparisons (right panel). In both cases the *All* distribution strongly first order stochastically dominates the *None* treatment ($p < 0.001$ for both statistics via Kolmogorov-Smirnov tests), suggesting that (i) subjects choose far simpler strategies on average and (ii) are far more likely to choose algorithmically simpler rather than complex strategies when exposed to our hypothesized drivers of aversion than when not.

---

complex one; we then compute the likelihood of choosing the relatively simple strategy among those pairs for each subject. Below, we complement this analysis with conditional logit regressions (also pre-registered), with the unit of analysis being the strategy choice game as a whole, to provide further statistical evidence.

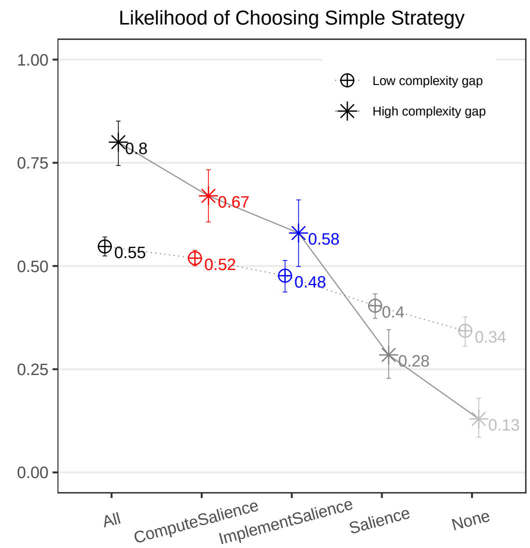

Figure 7: Likelihood of Choosing Simple Strategy by Complexity Gap *Notes: A subject's strategy choice in a game is analyzed as six pairwise comparisons. Complexity gap is the difference in complexity (i.e., the total number of transitions in the automaton representation) between the two strategies being compared. High (Low) complexity gap means being larger (weakly smaller) than 4, i.e., the median. The short vertical lines denote 95 percent confidence intervals by bootstrapping.*

Finally, the substantial changes in aversion to algorithmically complex strategies we document in pairwise comparisons (in the right-hand panels of Figures 5 and 6) are sensitive to the magnitude of the *gap* in algorithmic complexity between strategies in exactly the manner we would expect under the hypothesis of complexity aversion. To investigate this, we examine the likelihood of choosing the relatively simple strategy as a function of the difference in algorithmic complexity (i.e., difference in the total number of transitions) of the two strategies being compared—a statistic we will refer to as the "complexity gap" between the two strategies. Figure 7 plots the likelihood of choosing the relatively simple strategy when the complexity gap between a pair of strategies is weakly smaller than the median (i.e., 4) and larger than the median, respectively. When the complexity gap is relatively small, the difference between the *All* and *None* treatment is relatively small too (0.34 in *None* to 0.55 in *All*). However, when the complexity gap is large, the treatment differences become substantially larger, with the likelihood of choosing the relatively simpler strategy jumping from 0.13 in *None* to 0.80 in *All* – a more than six fold difference. This suggests that the treatment effects we document are dependent on the "dose" of algorithmic complexity in exactly the way we should expect. Figure 14 in Appendix B presents a finer-grained analysis, plotting the likelihood of choosing the relatively simple strategy at each level of the complexity gap, which reinforces this finding.

Result 1. *Subjects show a marked aversion to complex strategies. When all three hypothesized drivers of complexity aversion are present (the All treatment), subjects systematically choose algorithmically simpler strategies than when they are absent (the None treatment). This effect intensifies as the gap in algorithmic complexity between strategies grows larger.*

4.2 Three Faces of Complexity

In the previous subsection, we found that subjects show a systematic (and sizable) aversion to algorithmically complex strategies but only when the drivers of complexity aversion hypothesized in Section 2 are present. Next, we consider the effects of these three hypothesized drivers individually and how they each contribute to this overall effect.

In order to measure the effect of *implementation complexity*, we compare treatment *ComputeSalience* (which removes implementation complexity while retaining the two other facets of complexity) to treatment *All*. Returning to Figure 5, it is clear that (i) the average algorithmic complexity of chosen strategies is higher in *ComputeSalience* than *All* and (ii) the likelihood of choosing simpler strategies in pairwise comparisons is smaller (Wilcoxon rank sum test, $p < 0.01$ for both measures). Likewise, as is clear from Figure 6, removing implementation complexity leads to corresponding distributional changes in both measures (Kolmogorov-Smirnov test, $p < 0.01$).

Similarly, to measure the effect of *computational complexity*, we compare treatment *ImplementSalience* (which removes computational complexity while retaining implementation complexity and salience) to treatment *All*. Once again, relative to *All*, Figure 5 shows that (i) the average algorithmic complexity of chosen strategies is higher in *ImplementSalience* and (ii) the likelihood of choosing simpler strategies in pairwise comparisons is smaller ($p < 0.01$). Figure 6 reinforces this by showing that removing computational complexity leads to corresponding shifts in the distributions of both measures ($p < 0.01$).

Comparing treatment *Salience* (which removes both implementation and computational complexity) to treatment *All* allows us to measure the *joint effect* of implementation and computational complexity. As both Figures 5 and 6 show, these two drivers of complexity are not pure substitutes: the two drivers reinforce one another to some extent, producing a combined effect (on both average algorithmic complexity of chosen strategies and the likelihood of choosing simpler strategies in pairwise comparisons) that is significantly stronger than either driver has in isolation ($p < 0.01$ for all relevant comparisons).[45]

Finally, in order to measure the effect of the *salience* of simple strategies (our third hypothesized driver of complexity aversion), we compare treatment *Salience* (which removes implementation and computational complexity but retains salience) to treatment *None* (which removes all three drivers). Again, this has a significant effect ($p < 0.01$) on both measures, in terms of both means (Figure 5) and distributions (Figure 6).

Table 5: Impacts of Complexity and Strategic Uncertainty on Strategy Choice

	Strategy Choice (Conditional Logit)		
	(1)	(2)	(3)
complexity	0.306***	0.253***	0.196***
	(0.034)	(0.033)	(0.036)
Salience×complexity	-0.191***	-0.175***	-0.142***
	(0.039)	(0.036)	(0.042)
Computation×complexity	-0.174***	-0.166***	-0.069***
	(0.022)	(0.021)	(0.026)
Implementation×complexity	-0.135***	-0.128***	-0.147***
	(0.022)	(0.021)	(0.025)
BOA		0.619***	1.272***
		(0.168)	(0.312)
Salience×BOA			-0.238
			(0.448)
Computation×BOA			-2.186***
			(0.417)
Implementation×BOA			0.312
			(0.424)
No. of subjects	500	500	500

*Notes: Conditional logit regression with the unit of analysis being a strategy choice game as a whole. The dependent variable is an indicator of whether a strategy is chosen. The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. BOA denotes the size of the basin of attraction of a strategy, see Section 2.2. Salience, computation, and implementation are treatment indicators for salience, computational, and implementation complexity, respectively. Clustered standard errors at the subject level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.*

In Table 5 we provide further statistical support for these findings by estimating a series of conditional logit regressions (as outlined in our pre-registration). Here, the unit of analysis is a strategy choice game as a whole: in each game, a subject

[45] The fact that both implementation and computational complexity are (unavoidably) measured in the presence of salience, likewise suggests that neither (i) implementation complexity and salience nor (ii) computational complexity and salience are pure substitutes for one another but instead have a cumulative effect on behavior.

faces a set of seven strategies and chooses one. In specification (1), we regress an indicator for whether a strategy is chosen on its complexity (i.e., total number of transitions in its automaton representation), including interactions with treatment indicators for the presence of salience, computational, and implementation complexity. Results from (1) reinforce the preceding findings. In the *None* treatment (in which drivers of complexity aversion have been removed), subjects are drawn to algorithmically complex strategies (perhaps because of slight residual payoff dominance of more complex strategies in our design, see Table 4). However, introducing any of the sources of aversion produces a strong effect in the opposite direction: the interactions of each treatment indicator with strategy complexity are all negative and highly significant ($p < 0.01$), confirming that salience, computational, and implementation complexity each independently generate significant aversion to complex strategies. In specifications (2) and (3) (which we discuss in more detail in Section 4.4 below), we add controls for the size of the basin of attraction (BOA) and continue to find that each facet of complexity significantly reduces subjects' likelihood of choosing algorithmically complex strategies.[46]

Result 2. *Each of the three facets of complexity, individually, significantly reduces subjects' likelihood of choosing algorithmically complex strategies, contributing to the overall effect of complexity on strategic choice.*

Our design also allows us to explore the interaction between drivers of complexity aversion. As shown above, comparing *Salience* to *All*, we find that there is some degree of accumulation of the effects of computational and implementation complexity. However, our design allows us to go further because (as discussed in Section 3), our four treatments that include salience (*All*, *ImplementSalience*, *ComputeSalience*, and *Salience*) form a 2×2 factorial design crossing implementation and computational complexity, allowing us to study not only the main effects of each dimension but also their interaction.[47]

Examining Figure 5 and focusing on either statistic (complexity level of chosen strategies or likelihood of choosing the relatively simple strategy), the effect of adding

[46]Table 7 in Appendix B presents another specification with more controls for payoff characteristics of each strategy, including the diagonal payoff (i.e., the payoff from coordinating on the same strategy), the minimum payoff, and the maximum payoff across all opponent strategies. The treatment-specific complexity aversion effects remain significant and of similar magnitude.

[47]Of course, we cannot study all interactions directly. In particular, we cannot study implementation and computational complexity without describing the strategies to subjects, which means we cannot directly measure interactions between salience and either of these notions of complexity.

implementation (computational) complexity is larger when the other dimension is absent than when it is present. Specifically, the difference between *Salience* and *ImplementSalience* (*ComputeSalience*) treatments, which captures the effect of adding implementation (computational) complexity in isolation, is larger than the difference between the *ComputeSalience* (*ImplementSalience*) and the *All* treatments, which captures the effect of adding implementation (computational) complexity when the other dimension is already present.[48] This suggests that there is a *negative interaction* between the two dimensions of complexity: when one source of complexity is already present, the marginal effect of introducing the other is diminished. In other words, while each dimension independently generates aversion to complex strategies, their combined effect is sub-additive. Nonetheless, this sub-additivity does not undermine the overall monotonic pattern of treatment effects shown earlier: despite the diminishing marginal effect, adding a second dimension of complexity still significantly increases complexity aversion beyond what a single dimension produces. That is, the main effects of each dimension dominate the negative interaction, so the joint effect of implementation and computational complexity together remains substantially larger than the effect of either dimension alone.[49]

4.3 Payoff Consequences

How economically significant is the impact of complexity on strategy choice? One way to judge this is to study the severity of the payoff consequences of the complexity aversion we've documented. Selection of simpler strategies need not lead to lower payoffs in general, but they tend to do so in our experiment, by design.[50] This was a deliberate feature of our experiment that we included in the design in order to give us scope to assess the *strength* of complexity aversion relative to competing economic

[48] According to Figure 5, the difference in the average complexity of chosen strategies is 1.67 (2.07) between *Salience* and *ImplementSalience* (*ComputeSalience*) but 0.84 (1.24) between *ComputeSalience* (*ImplementSalience*) and *All*. The difference in the likelihood of choosing the relatively simple strategy is -0.16 (-0.22) between *Salience* and *ImplementSalience* (*ComputeSalience*) but -0.08 (-0.14) between *ComputeSalience* (*ImplementSalience*) and *All*.

[49] Table 8 in Appendix B provides further statistical support for this finding. We consider a conditional logit specification with a three-way interaction term, *computation*× *implementation*× *complexity*, and find its coefficient is significantly positive, confirming that the marginal complexity aversion generated by one dimension is attenuated in the presence of the other.

[50] This is due (i) to the fact that we selected strategies in which payoffs tend to increase with strategy complexity and (ii) perhaps more importantly, to the fact (discussed in more depth in Section 4.4 below) that less complex strategies in our design tend to be, on average, more strategically risky.

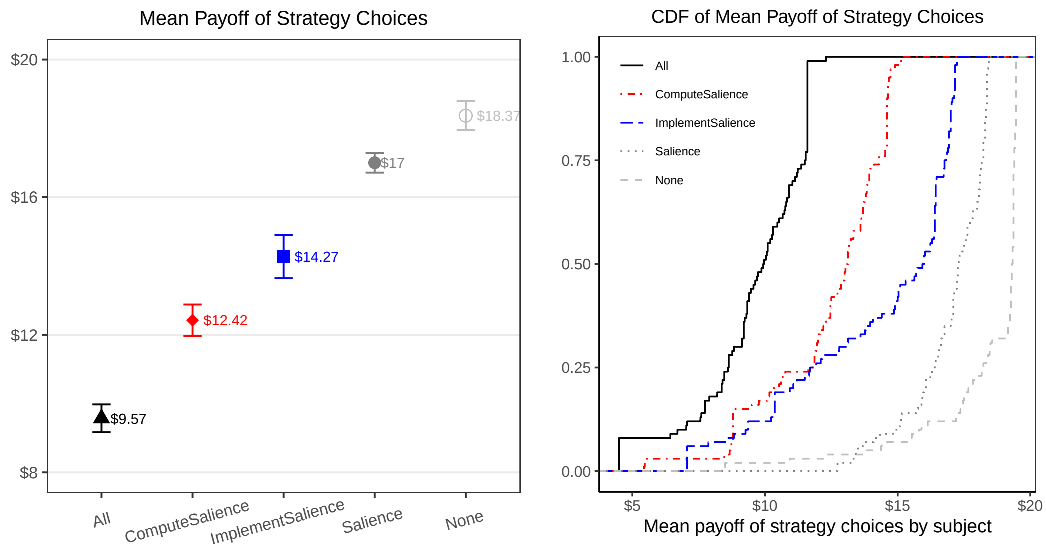

Figure 8: Mean payoff of Strategy Choices *Notes: Payoffs are in dollars. One dollar is two experiment tokens. The calculation focuses only on payoffs from the strategy choice stage and does not include potential penalties for mistakes from the Implementation Stage in All and ImplementSalience treatments. For the left panel, we first compute the average payoff of strategy choices for each subject and then calculate the treatment-level averages. The short vertical lines in the left panel denote 95 percent confidence intervals.*

motivations. Intuitively, to the degree subjects are willing to sacrifice a significant portion of their expected game payoffs simply to avoid complex strategies, we have evidence that the forces driving complexity aversion in this setting are behaviorally significant in a revealed preference sense. By making selection of simple strategies costly in terms of expected earnings, our experiment allows us to gather at least a lower bound estimate of the strength of complexity aversion in monetary terms.

The left panel of Figure 8 plots the dollar value of the mean number of experiment tokens earned in each treatment.[51] The right hand panel complements this by showing CDFs of subject-wise means of the same metric. The results reveal that subjects are willing to sacrifice a strikingly large proportion of their earnings to avoid complex strategies. When none of the drivers of complexity aversion are present subjects earn, on average, \$18.37. However when all of the drivers are present they earn only about half of this amount (\$9.57).[52] This is a substantial change given that the maximum payoff from strategy choices is \$20 (40 tokens).[53] Indeed, these results suggest that

[51] In order to put the treatments on an even footing, for this calculation we include only the base payment from strategy choices – we do not include the (small) payoff losses from the Implementation Stage. Note that this is a conservative choice: if we had included them, these payoff effects would be larger still.

[52] This difference is statistically significant via both Wilcoxon rank sum tests and Kolmogorov-Smirnov tests ($p < 0.001$).

[53] This is also sizable compared to payoffs from other repeated game experiments (e.g., in Fuden-

subjects are willing to sacrifice at least 44% of their earnings to avoid all three facets of complexity. What's more, these payoff consequences are contributed to by all three facets of complexity: *Salience* earnings are significantly lower than *None* (showing substantial payoff effects of salience), *ImplementSalience* earnings are significantly lower than *Salience* (showing payoff effects from implementation complexity) and *ComputeSalience* is significantly lower than *Salience* (showing payoff effects from computational complexity).[54] Thus all three faces of complexity are strong enough to induce subjects to leave a significant amount of money "on the table."

Result 3. *The influence of complexity is strong enough to induce subjects to sacrifice, on average, 44% of their potential monetary payoffs (over $8) when making their strategy choices. Each of the three drivers of complexity aversion we consider contributes significantly to this effect.*

4.4 Complexity versus Strategic Uncertainty

Another natural way to assess the economic significance of these findings is to compare complexity effects to the most important alternative drivers of strategy choice identified in the prior literature. As discussed in Section 2.2, the most significant driver of strategy choice documented so far in the experimental literature is *strategic uncertainty* as measured by the size of the basin of attraction (BOA) of strategies—a measure of the strategic risk players expose themselves to by employing a strategy. Returning to the regressions in Table 5, we can confirm that (as in the previous literature) strategic uncertainty is a major driver of behavior. Specification (2) includes the BOA of strategies and specification (3) interacts this with treatment variables related to complexity. The results reveal that although complexity continues to have strong independent impacts on strategy choice, the BOA has a significant role as well and in the predicted direction.[55] Indeed, these results suggest that complexity and BOA jointly shape strategy choice.

berg et al. (2012), subjects' payoffs range from $14 to $36 (including a $10 show-up fee); in Dal Bó & Fréchette (2019), subjects' payoffs range from $12.26 to $43.45).

[54]All pairwise differences are statistically significant via both Wilcoxon rank sum tests and Kolmogorov-Smirnov tests ($p < 0.001$).

[55]The BOA only interacts with one of our complexity drivers: the interaction between computation and BOA is large and significantly negative, indicating that computational complexity substantially reduces subjects' responsiveness to BOA. That is, unsurprisingly, when subjects must infer or compute payoff consequences themselves, they are less influenced by the BOA (strategic risk) of strategies in their choices. By contrast, the interactions of salience and implementation with BOA are not individually significant.

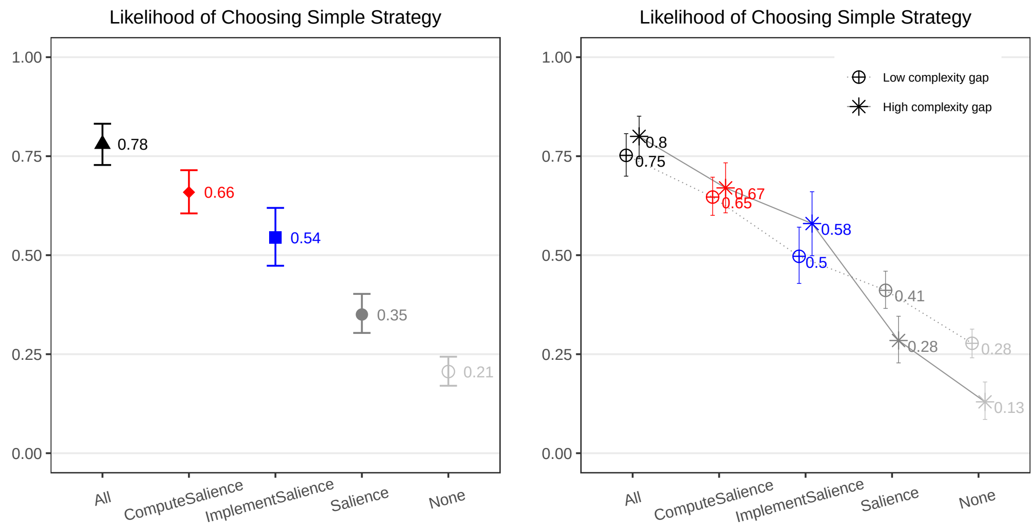

Figure 9: Likelihood of Choosing Simple Strategy – Complex Strategy Having Less Strategic Risk *Notes: Analysis is on only the pairwise comparisons in which BOA predicts that the complex strategy is more likely to be chosen. The left panel pools all these pairwise comparisons, while the right panel separately examines pairs with (weakly) lower-than-median complexity gaps (the median is 4) and those with higher-than-median complexity gaps. Complexity gap is the difference in complexity (i.e., the total number of transitions in the automaton representation) between the two strategies being compared. The short vertical lines denote 95 percent confidence intervals by bootstrapping.*

Theoretically, complexity and strategic uncertainty need not be related and, indeed, can be orthogonal to one another. In our experiment, however, we deliberately selected a set of strategies that often produce a direct tension between them (discussed in Section 3.1). In particular, our strategy choice game creates a number of cases (pairwise comparisons) in which the less strategically risky strategy is also more complex. In these cases, complexity aversion and aversion to strategic risk conflict, creating head-to-head contests between the two putative drivers of strategy choice. Because of this, we can assess the relative strength of complexity aversion and strategic risk avoidance simply by examining whether subjects choose complex or simple strategies more frequently in cases in which the two conflict.

The left panel of Figure 9 plots the likelihood of choosing the relatively simpler strategy in this subset of cases in which strategic risk and complexity aversion conflict. In the *None* treatment, the likelihood is only 0.21 (simpler strategies are selected 21% of the time), well below 0.5 (bootstrap-based test, $p < 0.001$). This is expected given that strategic uncertainty *should* be the primary driver of choice in *None*, where all three drivers of complexity aversion are removed. The likelihood increases to 0.35 in the *Salience* treatment, suggesting that one face of complexity—salience—begins to erode the influence of strategic uncertainty but strategic uncertainty nonetheless continues to dominate. However, adding either or both of the remaining drivers of complexity aversion causes complexity to dominate: the likelihood increases markedly to 0.54, 0.66, and 0.78 in the *ImplementSalience*, *ComputeSalience*, and *All* treatments, respectively. All of these likelihoods are above 0.5, indicating that the influence of complexity tends to (on net) dominate that of strategic uncertainty when computational and/or implementation complexity is present.[56] Moreover, at the subject level, in cases where complexity and strategic uncertainty conflict, we find that 57%, 65%, and 87% of subjects in the *ImplementSalience*, *ComputeSalience*, and *All* treatments are individually strictly more likely to choose relatively simple strategies (i.e., the individual-level likelihood is strictly larger than 0.5) while only 10% and 27% of subjects do so in the *None* and *Salience* treatments.

The right panel of Figure 9 breaks this down by cases in which the complexity gap is above or weakly below the median (of 4). As the complexity gap increases, the likelihood of choosing the relatively simple strategy (rather than the less strategically risky strategy) increases in treatments with computational and/or implementation complexity (while it decreases in the *Salience* and *None* treatments). When focusing on cases with a higher-than-median complexity gap, the increasing dominance of complexity is even starker: across treatments (from *None* to *All*), the likelihoods are 0.13, 0.28, 0.58, 0.67, and 0.80, with the latter three being substantially higher than 0.5.[57] Again, these findings suggest that the influence of computational and implementation complexity on strategy choice (when combined with salience effects) is comparable to and in fact somewhat stronger than that of strategic uncertainty.

Result 4. *When complexity and strategic uncertainty have opposite predictions on strategy choice, the impact of computational and implementation complexity dominates the influence of strategic uncertainty.*

4.5 Robustness

Our experimental design allows us to assess the robustness of these results in several ways and to thereby deepen our interpretation of our findings.

[56]The likelihood in *All*, *ComputeSalience*, and *ImplementSalience* treatments is higher in cases where complexity and strategic uncertainty conflict compared to that in the full sample. This is because, by design, in a large proportion of these cases, the complexity gap between strategy pairs is relatively high.

[57]Under the bootstrap-based test against a null of 0.5, *ImplementSalience* has $p = 0.058$ while each of the other four treatments has $p < 0.001$.

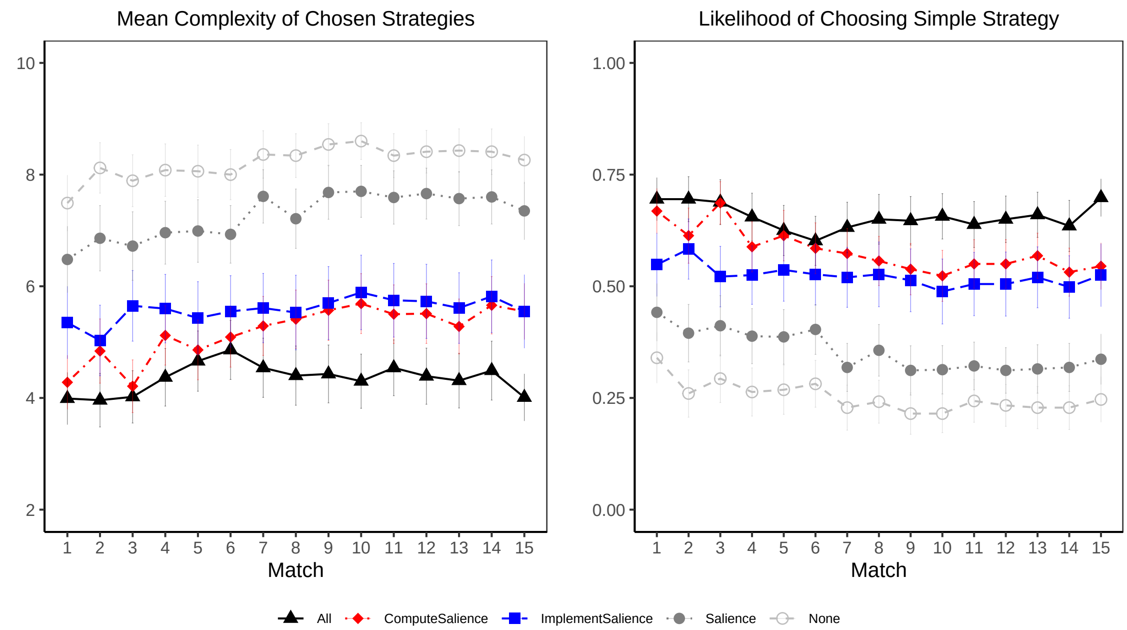

Figure 10: Complexity Level of Chosen Strategies and Likelihood of Choosing Simple Strategy – Across Match *Notes: The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. When computing the likelihood of choosing (the relatively) simple strategy in the right panel, a subject's strategy choice in a game is analyzed as six pairwise comparisons. The short vertical lines denote 95 percent confidence intervals.*

4.5.1 Learning & Experience

As in most contemporary repeated games experiments, subjects in our experiment play the same supergame multiple times (fifteen in total) and receive feedback after each match, creating scope for learning. This feature of our design allows us to examine whether the complexity aversion we've measured is transient or if it is instead a stable influence on experienced subjects' strategy choices.

Figure 10 plots the average complexity level of chosen strategies (left panel) and the likelihood of choosing the relatively simple strategy in pairwise comparisons (right panel) across the 15 supergames (matches) in our experiment. We make two observations. First, the treatment effects and their monotonic pattern—complexity aversion strengthens progressively from *None* to *Salience* to *ImplementSalience* and *Compute-Salience* to *All*—are remarkably robust, persisting across all 15 supergames. Second, the magnitude of these treatment effects (especially of computational and implementation complexity) does not diminish with experience; if anything, it slightly increases in later matches, suggesting that learning does not weaken complexity aversion but rather seems to reinforce it. Taken together, these results suggest that the complexity effects we document reflect a persistent feature of how subjects evaluate and choose strategies rather than (for instance) a transient artifact of unfamiliarity with the task. To further confirm this, we replicate our main regression analysis using only the last five matches and obtain results that are both qualitatively and quantitatively consistent with those from the full sample (see Table 11 in Appendix C.1).

Result 5. *Aversion to complexity does not diminish with learning or experience, suggesting that the influence of complexity on strategy choices is systematic and persistent.*

4.5.2 Cognitive Ability

A recent literature (e.g., Proto et al. 2019, 2022) studies the effects of cognitive ability on behavior in repeated games. We explore whether the treatment effects we document are heterogeneous across subjects with different levels of cognitive ability.

After the main part of the experiment (15 supergames), subjects are asked to answer a set of ten cognitive questions. We split subjects into two groups based on whether their cognitive test score exceeds the median of 6, and examine whether treatment effects differ between high- and low-cognitive ability subjects. Figure 11 plots our two main statistics on strategy choices separately for the two groups. We find that the overall complexity effect does not diminish but, if anything, becomes slightly stronger among subjects with above-median cognitive scores. In particular, in terms of either the average complexity of chosen strategies or the likelihood of choosing the relatively simple strategy, the difference between the *All* and *None* treatments is slightly larger among high-cognitive ability subjects.[58] These findings suggest that complexity aversion plausibly reflects a deliberate response to the cognitive costs of complex strategies rather than an inability to process them, confusion, or inattentiveness: high-cognitive ability subjects, who are potentially better at recognizing the costs that complexity imposes, respond to these costs at least as strongly as (or even more strongly than) low-cognitive ability subjects.

The regression results presented in Table 10 in Appendix B reveal a more nuanced picture of how cognitive ability interacts with the different dimensions of complexity. Among high-cognitive ability subjects, salience and computational complexity generate relatively stronger complexity aversion (i.e., the coefficient on the treatment indi-

[58] Focusing on pairwise comparisons in which the complexity gap is above median, the likelihood of choosing the relatively simple strategy is 0.85 vs. 0.07 in *All* and *None* treatments among high-cognitive ability subjects, while 0.77 vs. 0.17 among low-cognitive ability subjects.

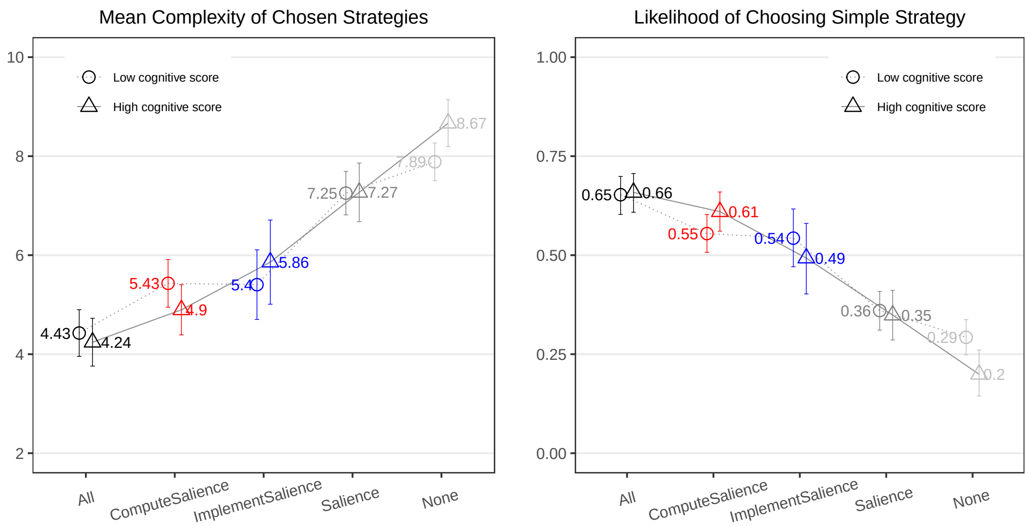

Figure 11: Complexity Level of Chosen Strategies and Likelihood of Choosing Simple Strategy – By Cognitive Score *Notes: The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. When computing the likelihood of choosing (the relatively) simple strategy in the right panel, a subject's strategy choice in a game is analyzed as six pairwise comparisons. High (Low) cognitive score means being higher (weakly lower) than the median of 6. The short vertical lines denote 95 percent confidence intervals by bootstrapping.*

cator and complexity interaction term is larger in magnitude), while implementation complexity has a relatively weaker effect. One interpretation is that high-cognitive ability subjects are better at recognizing the strategic value of simplicity as a coordination device, generating a salience effect that is intensified by computational complexity; by contrast, the cognitive burden of round-by-round implementation is plausibly less costly for high-cognitive ability subjects.[59]

Result 6. *The overall effect of complexity on strategy choice does not diminish among subjects with higher cognitive ability, suggesting that complexity aversion likely isn't driven primarily by confusion, inattention or mistakes. High-cognitive ability subjects show stronger responses to salience and computational complexity but relatively weaker responses to implementation complexity.*

4.5.3 Comprehension

Table 9 in Appendix B replicates our main regression analysis, considering only the subjects (453 out of 500) who answered all the comprehension questions on instruc-

[59]The regression results also show that high-cognitive ability subjects are more responsive to BOA in the *None* and *Salience* treatments, where strategic uncertainty is expected to be pronounced, than low-cognitive ability subjects. This is broadly consistent with Hanaki et al. (2016), who find in 2×2 coordination games that higher cognitive ability subjects are more sensitive to the presence of strategic uncertainty.

tions correctly on their first try. The results are both qualitatively and quantitatively consistent with those from the full sample. This suggests that our results persist among subjects who are most attentive, highest effort, and least likely to be confused about the rules of the experiment.

4.5.4 Laboratory Experiment

As we say in Section 3.3, we conducted laboratory sessions in which subjects played the strategy choice game simultaneously with randomly paired partners. Appendix C.2 presents the results from these lab sessions (94 subjects in total). Although our sample sizes are very small, we find that our main findings replicate in the lab: the average complexity of chosen strategies decreases monotonically as more dimensions of complexity are present, dropping from 8.05 in *None* to 4.07 in *All*, and the likelihood of choosing the relatively simple strategy increases from 0.28 to 0.7 from *None* to *All* (Figure 18). Thus, we find similar overall evidence that complexity significantly influences strategic choice. Our conditional logit regressions on this sample (Table 12) confirm that both computational and implementation complexity generate significant aversion to complex strategies (with coefficients that are in fact somewhat larger in magnitude than in the Prolific sample). Conversely, the main notable difference we find is that the effect of salience alone is smaller and is not statistically significant among lab participants.

5 Conclusion

Our results suggest that complexity shapes strategy choice in repeated games in three fundamental ways. First, we find that subjects exhibit a strong aversion to implementing algorithmically complex strategies: when required to execute their chosen strategy themselves, they systematically shift toward simpler strategies. Second, the computational complexity of inferring the payoff consequences of strategies leads subjects to systematically avoid algorithmically complex strategies. Finally, we find that subjects use the algorithmic complexity of strategies as a *coordination device*, using the salience of simple strategies to manage strategic risk in repeated games. These effects are large, costly and comparable in strength to the well-documented first-order effects (Dal Bó & Fréchette 2011, 2018, Blonski et al. 2011, Blonski & Spagnolo 2015, Calford & Oprea 2017, Embrey, Fréchette & Yuksel 2017, Vespa & Wilson 2017, Boczoń et al. 2024) strategic risk has in shaping strategy selection. We show that this aversion to complex strategies is robust to learning and variation in measured cognitive ability and occurs both in online and lab experiments.

In order to measure these effects, we study an experiment in which subjects directly choose their strategies from a menu rather than directly choosing actions, round-by-round, in the game. This feature of the design is necessary for the kind of control over the strategy menu we need in order to cleanly measure the effects of complexity (given the infinite size of the strategy space available in direct-action choice games). Nonetheless, it is natural to wonder how relevant our finding of a severe bias towards simple strategies is to more naturalistic direct-action versions of the task.

We think there is a strong case that the complexity effects we measure can be expected to similarly shape strategy choice in more naturalistic settings. For instance, our finding of implementation complexity is derived from asking subjects to implement exactly the same action sequences they would be required to implement in a direct action game.[60] Likewise, the computational complexity of the task of forecasting the payoff consequences of strategies is nearly identical in our task and in direct action games. Finally, we think that the salience of simple strategies should be expected to be at least as operative in more naturalistic games as it is in our explicit strategy choice games. Indeed, we suspect that the main artificiality introduced by our design is that it may make the availability of more complex strategies *more* salient relative to simpler strategies than they would be in direct action games. After all, were they not explicitly made available in our menu, subjects may not have otherwise even conceived of elaborate strategies like D-TF2T or D-TF3T, pushing subjects even further towards the use of simple strategies than in our data.[61] Given this, we think that there is a case that our results may even represent a lower bound estimate of the effects of complexity on strategy choice in more naturalistic versions of the game.

These results have both empirical and theoretical implications. Empirically, our results fill in a gap in our understanding of how subjects select strategies in repeated

[60]Indeed, in order to make this mapping even more direct, we inject realistic levels of noise into the implementation of counterpart strategies, estimated from prior direct choice data. Doing this makes the likelihood of off-path play—and therefore the subjective costs of automata states and transitions in the algorithmic structure of the strategy that are only reached off-path—as close as possible to the level observed in direct choice data.

[61]Indeed, this may point to a fourth "face" of complexity, not measured in our design. As Oprea (2026) emphasizes, in addition to the complexity of *implementing* rules, many tasks also generate a complexity of *identifying* high quality rules. Indeed the subjective costs of identifying behavioral rules may even, in some settings, be higher than the costs of implementing them.

games. An important (but thus far unexplained) regularity in experimental studies of strategy selection is that subjects overwhelmingly tend to confine their strategic choices to a relatively small menu of algorithmically simple strategies (Dal Bó & Fréchette 2018). Our results provide a direct explanation for this regularity: subjects dislike executing complex strategies and coordinate using simple strategies, both of which shrink the set of strategies subjects use in games. Our findings suggest that subjects use complexity to refine the infinite space of strategies to a small consideration set, and use strategic risk to choose between the surviving strategy options. This account seems to unify and is highly consistent with prior findings from the long experimental literature on behavior in repeated and dynamic games.

Theoretically, our results support and extend the findings from an important literature that uses complexity as a tool for refining the equilibrium set in repeated games. The automata literature (Aumann 1981, Neyman 1985, Rubinstein 1986, Abreu & Rubinstein 1988, Banks & Sundaram 1990) shows that if agents have even infinitesimal aversion to implementing algorithmically complex strategies, the notoriously large set of strategies supportable as equilibria in dynamic settings shrinks significantly. In this sense, acknowledging complexity in theoretical models eases the indeterminacy of the Folk theorem, making repeated game theory more predictive. Our findings strongly support the premises of these models by showing that subjects do indeed attach implicit costs to algorithmically complex strategies, and these costs scale with the size of the automaton representation as the theory assumes. Moreover, our results suggest that complexity aversion can arise through three distinct channels—implementation, computation, and salience. Incorporating all these channels may therefore offer a more complete account of how complexity shapes strategic behavior. Our findings also suggest that complexity costs are likely considerably larger than those usually hypothesized in automata models. Because of this, our results imply that these models can be made substantially more restrictive, improving their capacity to refine the equilibrium set, and yielding sharper predictions about behavior in repeated games. Exploring the theoretical consequences of both the multidimensional nature and the much larger magnitude of complexity costs implied by our findings, and testing the refined predictions in further experiments, seems a highly valuable next step for this literature.

References

Abeler, J. & Jäger, S. (2015), 'Complex tax incentives', *American Economic Journal: Economic Policy* **7**(3), 1–28.

Abreu, D. & Rubinstein, A. (1988), 'The structure of nash equilibrium in repeated games with finite automata', *Econometrica* **56**(6), 1259–1281.

Aumann, R. J. (1981), 'Survey of repeated games', *Essays in game theory and mathematical economics in honor of Oskar Morgenstern* .

Ba, C., Bohren, J. A. & Imas, A. (2025), Over- and underreaction to information: Belief updating with cognitive constraints, working paper.

Banks, J. S. & Sundaram, R. K. (1990), 'Repeated games, finite automata, and complexity', *Games and Economic Behavior* **2**(2), 97–117.

Banovetz, J. & Oprea, R. (2023), 'Complexity and procedural choice', *American Economic Journal: Microeconomics* **15**(2), 384–413.

Ben-porath, E. (1990), 'The complexity of computing a best response automaton in repeated games with mixed strategies', *Games and Economic Behavior* **2**(1), 1–12.

Bernheim, B. D. & Sprenger, C. (2020), 'On the empirical validity of cumulative prospect theory: Experimental evidence of rank-independent probability weighting', *Econometrica* **88**(4), 1363–1409.

Blonski, M., Ockenfels, P. & Spagnolo, G. (2011), 'Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence', *American Economic Journal: Microeconomics* **3**(3), 164–92.

Blonski, M. & Spagnolo, G. (2015), 'Prisoners' other Dilemma', *Int J Game Theory* **44**, 61–81.

Boczoń, M., Vespa, E., Weidman, T. & Wilson, A. J. (2024), 'Testing models of strategic uncertainty: Equilibrium selection in repeated games', *Journal of the European Economic Association* **23**(2), 784–814.

Bordalo, P., Gennaioli, N. & Shleifer, A. (2022), 'Salience', *Annual Review of Economics* **14**(Volume 14, 2022), 521–544.

Calford, E. & Oprea, R. (2017), 'Continuity, inertia and strategic uncertainty: A test of the theory of continuous time games', *Econometrica* **85**(3), 915–935.

Cason, T. N. & Mui, V.-L. (2019), 'Individual versus group choices of repeated game strategies: A strategy method approach', *Games and Economic Behavior* **114**, 128–145.

Chater, N. (1999), 'The search for simplicity: A fundamental cognitive principle?', *The Quarterly Journal of Experimental Psychology Section A* **52**(2), 273–302.

Chater, N. & Vitányi, P. (2003), 'Simplicity: a unifying principle in cognitive science?', *Trends in Cognitive Sciences* **7**(1), 19–22.

Chatterjee, K. & Sabourian, H. (1998), N-person bargaining and strategic complexity, working paper.

Chatterjee, K. & Sabourian, H. (2000), 'Multiperson bargaining and strategic complexity', *Econometrica* **68**(6), 1491–1509.

Chatterjee, K. & Sabourian, H. (2009), *Game Theory and Strategic Complexity*, Springer New York, New York, NY, pp. 4098–4114.

Chen, D. L., Schonger, M. & Wickens, C. (2016), 'otree – an open-source platform for laboratory, online, and field experiments', *Journal of Behavioral and Experimental Finance* **9**, 88–97.

Dal Bó, P. & Fréchette, G. R. (2011), 'The evolution of cooperation in infinitely repeated games: Experimental evidence', *American Economic Review* **101**(1), 411–29.

Dal Bó, P. & Fréchette, G. R. (2018), 'On the determinants of cooperation in infinitely repeated games: A survey', *Journal of Economic Literature* **56**(1), 60–114.

Dal Bó, P. & Fréchette, G. R. (2019), 'Strategy choice in the infinitely repeated prisoner's dilemma', *American Economic Review* **109**(11), 3929–52.

Dal Bó, P., Fréchette, G. R. & Kim, J. (2021), 'The determinants of efficient behavior in coordination games', *Games and Economic Behavior* **130**(C), 352–368.

de Clippel, G., Moscariello, P., Ortoleva, P. & Rozen, K. (2025), Caution in the face of complexity, working paper.

Embrey, M., Fréchette, G. R. & Yuksel, S. (2017), 'Cooperation in the finitely repeated prisoners dilemma', *The Quarterly Journal of Economics* **133**(1), 509–551.

Embrey, M., Mengel, F. & Peeters, R. (2017), Eliciting strategies in indefinitely repeated games of strategic substitutes and complements, Working paper series, Department of Economics, University of Sussex Business School.

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

Enke, B., Graeber, T. & Oprea, R. (2025), 'Complexity and time', *Journal of the European Economic Association* **23**(5), 1838–1867.

Enke, B. & Shubatt, C. (2023), Quantifying lottery choice complexity, working paper.

Feldman, J. (2016), 'The simplicity principle in perception and cognition', *WIREs Cognitive Science* **7**(5), 330–340.

Fréchette, G. R. & Yuksel, S. (2017), 'Infinitely repeated games in the laboratory: four perspectives on discounting and random termination', *Experimental Economics* **20**(2), 279–308.

Friedman, J. W. (1971), 'A non-cooperative equilibrium for supergames', *The Review of Economic Studies* **38**(1), 1–12.

Fudenberg, D., Levine, D. & Maskin, E. (1994), 'The folk theorem with imperfect public information', *Econometrica* **62**(5), 997–1039.

Fudenberg, D. & Maskin, E. (1986), 'The folk theorem in repeated games with discounting or with incomplete information', *Econometrica* **54**(3), 533–554.

Fudenberg, D., Rand, D. G. & Dreber, A. (2012), 'Slow to anger and fast to forgive: Cooperation in an uncertain world', *American Economic Review* **102**(2), 720–49.

Gale, D. & Sabourian, H. (2005), 'Complexity and competition', *Econometrica* **73**(3), 739–769.

Gale, D. & Sabourian, H. (2006), 'Markov equilibria in dynamic matching and bargaining games', *Games and Economic Behavior* **54**(2), 336–352.

Gilboa, I. (1988), 'The complexity of computing best-response automata in repeated games', *Journal of Economic Theory* **45**(2), 342–352.

Gill, D. & Rosokha, Y. (2024), 'Beliefs, learning, and personality in the indefinitely repeated prisoner's dilemma', *American Economic Journal: Microeconomics* **16**(3), 25983.

Golman, R. & Page, S. E. (2010), 'Basins of attraction and equilibrium selection under different learning rules', *Journal of Evolutionary Economics* **20**, 49–72.

Guan, M., Oprea, R. & Yuksel, S. (2025), Beyond instrumental value: How complexity shapes information demand, working paper.

Hanaki, N., Jacquemet, N., Luchini, S. & Zylbersztejn, A. (2016), 'Cognitive ability and the effect of strategic uncertainty', *Theory and Decision* **81**(1), 101–121.

Harsanyi, J. & Selten, R. (1988), *A General Theory of Equilibrium Selection in Games*, Vol. 1, 1 edn, The MIT Press.

He, G. (2022), The easier the better? an experiment on strategy choice and complexity, Unpublished.

Jakobsen, A. M. (2020), 'A model of complex contracts', *American Economic Review* **110**(5), 124373.

Jones, M. T. (2014), 'Strategic complexity and cooperation: An experimental study', *Journal of Economic Behavior and Organization* **106**, 352–366.

Kalai, E. (1990), Bounded rationality and strategic complexity in repeated games, *in* T. Ichiishi, A. Neyman & Y. Tauman, eds, 'Game Theory and Applications', Economic Theory, Econometrics, and Mathematical Economics, Academic Press, San Diego, pp. 131–157.

Kalai, E. & Stanford, W. (1988), 'Finite rationality and interpersonal complexity in repeated games', *Econometrica* **56**(2), 397–410.

Murawski, C. & Bossaerts, P. (2017), 'Computational complexity and human decision-making', *Trends in Cognitive Sciences* **21**(12), 917–929.

Neyman, A. (1985), 'Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma', *Economics Letters* **19**(3), 227–229.

Oprea, R. (2020), 'What makes a rule complex?', *American Economic Review* **110**(12), 3913–51.

Oprea, R. (2024), 'Decisions under risk are decisions under complexity', *American Economic Review* **114**(12), 37893811.

Oprea, R. (2026), Complexity and its measurement, *in* 'Handbook of Experimental Methods in the Social Science'. Forthcoming.

Oprea, R. & Yuksel, S. (2022), 'Social exchange of motivated beliefs', *Journal of the European Economic Association* **20**(2), 667–699.

Papadimitriou, C. H. (1992), 'On players with a bounded number of states', *Games and Economic Behavior* **4**(1), 122–131.

Proto, E., Rustichini, A. & Sofianos, A. (2019), 'Intelligence, personality, and gains from cooperation in repeated interactions', *Journal of Political Economy* **127**(3), 1351–1390.

Proto, E., Rustichini, A. & Sofianos, A. (2022), 'Intelligence, Errors, and Cooperation in Repeated Interactions', *The Review of Economic Studies* **89**(5), 2723–2767.

Puri, I. (2025), 'Simplicity and risk', *The Journal of Finance* **80**(2), 1029–1080.

Robson, A. J. (2003), 'The evolution of rationality and the red queen', *Journal of Economic Theory* **111**(1), 1–22.

Romero, J. & Rosokha, Y. (2018), 'Constructing strategies in the indefinitely repeated prisoner's dilemma game', *European Economic Review* **104**, 185–219.

Romero, J. & Rosokha, Y. (2019), 'The evolution of cooperation: The role of costly strategy adjustments', *American Economic Journal: Microeconomics* **11**(1), 299–328.

Romero, J. & Rosokha, Y. (2023), 'Mixed strategies in the indefinitely repeated prisoner's dilemma', *Econometrica* **91**(6), 2295–2331.

Rubinstein, A. (1986), 'Finite automata play the repeated prisoner's dilemma', *Journal of Economic Theory* **39**(1), 83–96.

Sabourian, H. (2004), 'Bargaining and markets: complexity and the competitive outcome', *Journal of Economic Theory* **116**(2), 189–228.

Sandholm, W. H. (2010), *Population Games and Evolutionary Dynamics*, The MIT Press.

Schelling, T. C. (1960), *The Strategy of Conflict*, Cambridge: Harvard University Press.

Vespa, E. & Wilson, A. J. (2017), Experimenting with equilibrium selection in dynamic games, working paper.

Wilson, A. J. & Wu, H. (2017), 'At-will relationships: How an option to walk away affects cooperation and efficiency', *Games and Economic Behavior* **102**, 487–507.

ONLINE APPENDIX

A Studied Strategies and Games

Table 6: The Studied Strategies

Strategy	Complexity	Description
ALLD	1S-0T-a	Always choose A.
GRIM	2S-1T-a	Choose B until you see your counterpart chose A, after which switch to A forever.
TFT	2S-2T	In round 1 choose B. After round 1: choose whatever your counterpart chose in the previous round.
D-GRIM2	4S-3T-a	In round 1 choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.
D-TF2T	4S-4T	In round 1 choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.
D-GRIM3	5S-4T-a	In round 1 choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever.
D-TF3T	5S-5T	In round 1 choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B.

Notes: A and B denote the available action choices in the prisoner's dilemma shown in Table 1. The second column shows the algorithmic complexity of strategies in the automaton representation: xS-yT denotes that the number of states (s-complexity) of a strategy is x and the number of transitions in excess of the number of states (t-complexity) is y; a indicates the strategy has absorbing states (see Section 2.1).

B Additional Results

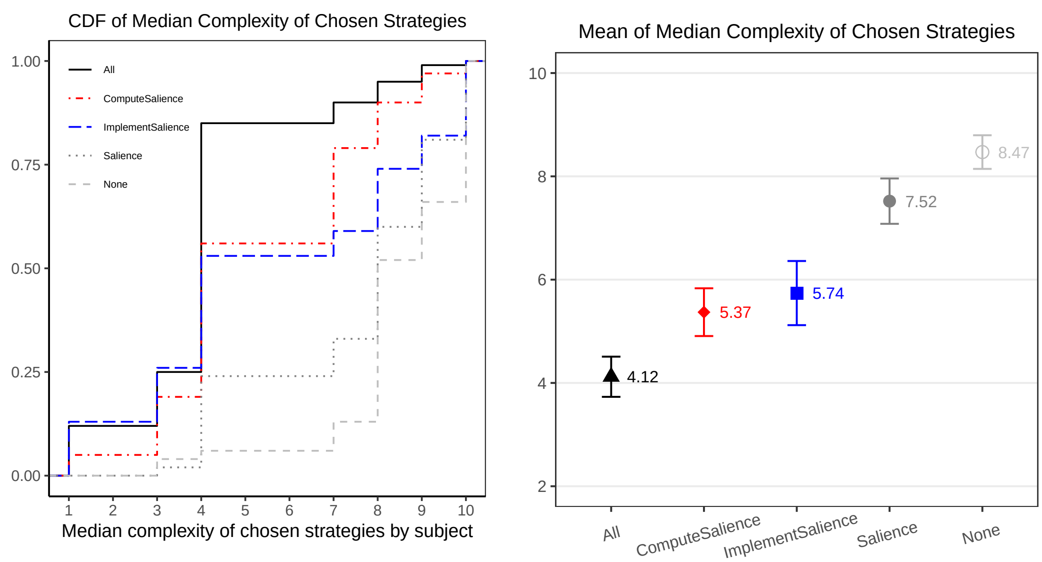

Figure 12: Median Complexity Level of Chosen Strategies *Notes: The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. For the right panel, we first compute the median complexity level of chosen strategies for each subject and then calculate the treatment-level averages. The short vertical lines in the right panel denote 95 percent confidence intervals.*

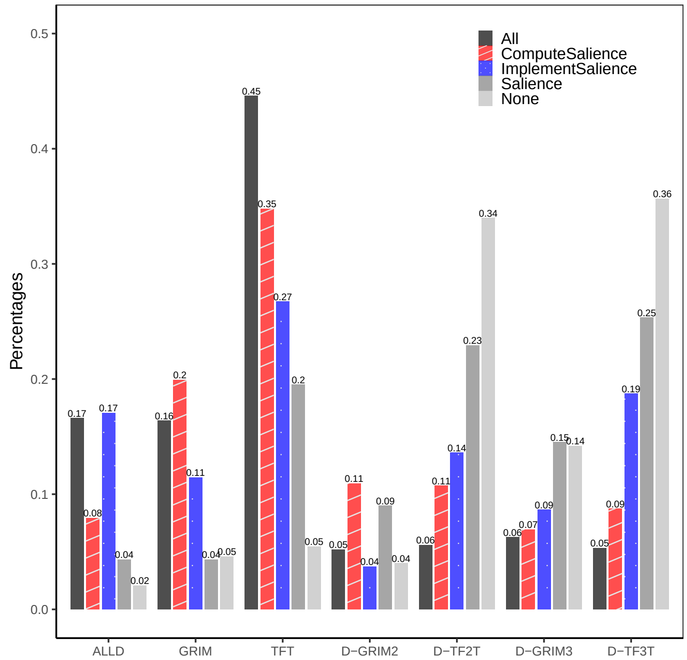

Figure 13: Distribution of Chosen Strategies Across Treatments

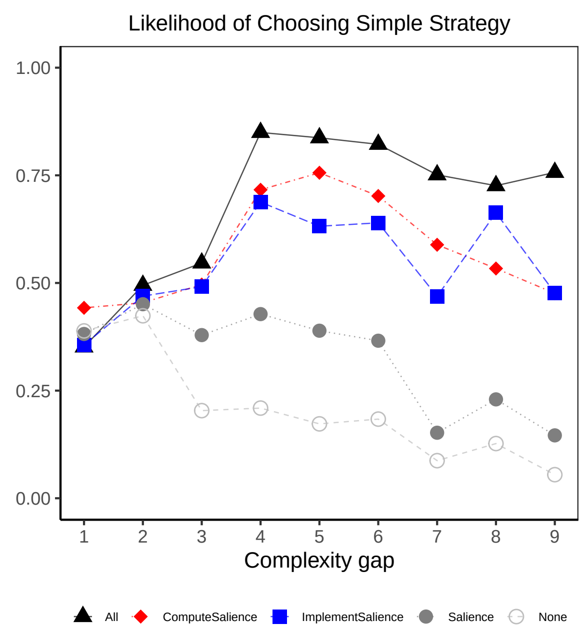

Figure 14: Likelihood of Choosing Simple Strategy Along Complexity Gap *Notes: A subject's strategy choice in a game is analyzed as six pairwise comparisons. Complexity gap is the difference in complexity (i.e., the total number of transitions in the automaton representation) between the two strategies being compared.*

Table 7: Impacts of Complexity and Strategic Uncertainty on Strategy Choice – More Controls

	Strategy Choice (Conditional Logit)
complexity	0.041
	(0.046)
Salience×complexity	-0.189***
	(0.057)
Computation×complexity	-0.119***
	(0.038)
Implementation×complexity	-0.225***
	(0.037)
BOA	1.754***
	(0.417)
Salience×BOA	0.065
	(0.516)
Computation×BOA	-1.555***
	(0.453)
Implementation×BOA	1.089**
	(0.471)
diagonal payoff	-0.008
	(0.013)
minimum payoff	-0.012
	(0.077)
maximum payoff	0.118***
	(0.022)
No. of subjects	500

*Notes: Conditional logit regression with the unit of analysis being a strategy choice game as a whole. The dependent variable is an indicator of whether a strategy is chosen. The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. BOA denotes the size of the basin of attraction of a strategy, see Section 2.2. Salience, computation, and implementation are treatment indicators for salience, computational, and implementation complexity, respectively. Three controls for payoff characteristics of each strategy, including the diagonal payoff (i.e., payoff from coordinating on the same strategy), the minimum payoff, and the maximum payoff across all opponent strategies. Clustered standard errors at the subject level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.*

Table 8: Interaction Between Computational and Implementation Complexity

	Strategy Choice (Conditional Logit)	
	(1)	(2)
complexity	0.086***	0.100***
	(0.026)	(0.027)
Computation×complexity	-0.119***	-0.149***
	(0.032)	(0.033)
Implementation×complexity	-0.202***	-0.233***
	(0.037)	(0.041)
Computation×Implementation×complexity	0.090**	0.156***
	(0.043)	(0.051)
BOA	0.891***	0.674**
	(0.312)	(0.339)
Computation×BOA	-2.115***	-1.476***
	(0.417)	(0.508)
Implementation×BOA	0.556	1.070*
	(0.430)	(0.560)
Computation×Implementation×BOA		-1.528*
		(0.870)
No. of subjects	400	400

*Notes: Analysis is on the four treatments with salience. Clustered standard errors at the subject level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.*

Table 9: Impacts of Complexity and Strategic Uncertainty on Strategy Choice – All Comprehension Questions Answered Correctly at First Try

	Strategy Choice (Conditional Logit)		
	(1)	(2)	(3)
complexity	0.310***	0.253***	0.187***
	(0.035)	(0.034)	(0.035)
Salience×complexity	-0.187***	-0.171***	-0.124***
	(0.041)	(0.037)	(0.043)
Computation×complexity	-0.187***	-0.178***	-0.077***
	(0.024)	(0.022)	(0.027)
Implementation×complexity	-0.133***	-0.126***	-0.151***
	(0.023)	(0.022)	(0.027)
BOA		0.648***	1.408***
		(0.176)	(0.312)
Salience×BOA			-0.424
			(0.459)
Computation×BOA			-2.283***
			(0.457)
Implementation×BOA			0.439
			(0.456)
No. of subjects	453	453	453

*Notes: Analysis is on only the subjects who answered all comprehension questions correctly at their first try (453 subjects). Clustered standard errors at the subject level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.*

Table 10: Impacts of Complexity and Strategic Uncertainty on Strategy Choice – By Cognitive Score

	Strategy Choice (Conditional Logit)					
	High Cognitive Score			Low Cognitive Score		
	(1)	(2)	(3)	(1)	(2)	(3)
complexity	0.428***	0.371***	0.249***	0.240***	0.191***	0.174***
	(0.079)	(0.078)	(0.081)	(0.033)	(0.033)	(0.037)
Salience×complexity	-0.308***	-0.284***	-0.209**	-0.126***	-0.116***	-0.108**
	(0.085)	(0.079)	(0.088)	(0.041)	(0.038)	(0.047)
Computation×complexity	-0.212***	-0.203***	-0.050	-0.147***	-0.140***	-0.079**
	(0.034)	(0.032)	(0.039)	(0.029)	(0.027)	(0.034)
Implementation×complexity	-0.112***	-0.107***	-0.106***	-0.152***	-0.144***	-0.177***
	(0.033)	(0.032)	(0.038)	(0.029)	(0.027)	(0.033)
BOA		0.526*	1.699***		0.644***	0.855**
		(0.275)	(0.553)		(0.212)	(0.375)
Salience×BOA			-0.326			-0.064
			(0.757)			(0.553)
Computation×BOA			-3.967***			-1.255**
			(0.774)			(0.500)
Implementation×BOA			-0.150			0.621
			(0.702)			(0.529)
No. of subjects	211	211	211	289	289	289

*Notes: Conditional logit regression with the unit of analysis being a strategy choice game as a whole. The dependent variable is an indicator of whether a strategy is chosen. High (Low) cognitive score means being higher (weakly lower) than the median score of 6. The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. BOA denotes the size of the basin of attraction of a strategy, see Section 2.2. Salience, computation, and implementation are treatment indicators for salience, computational, and implementation complexity, respectively. Clustered standard errors at the subject level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.*

C More on Robustness

C.1 Last Five Matches Only

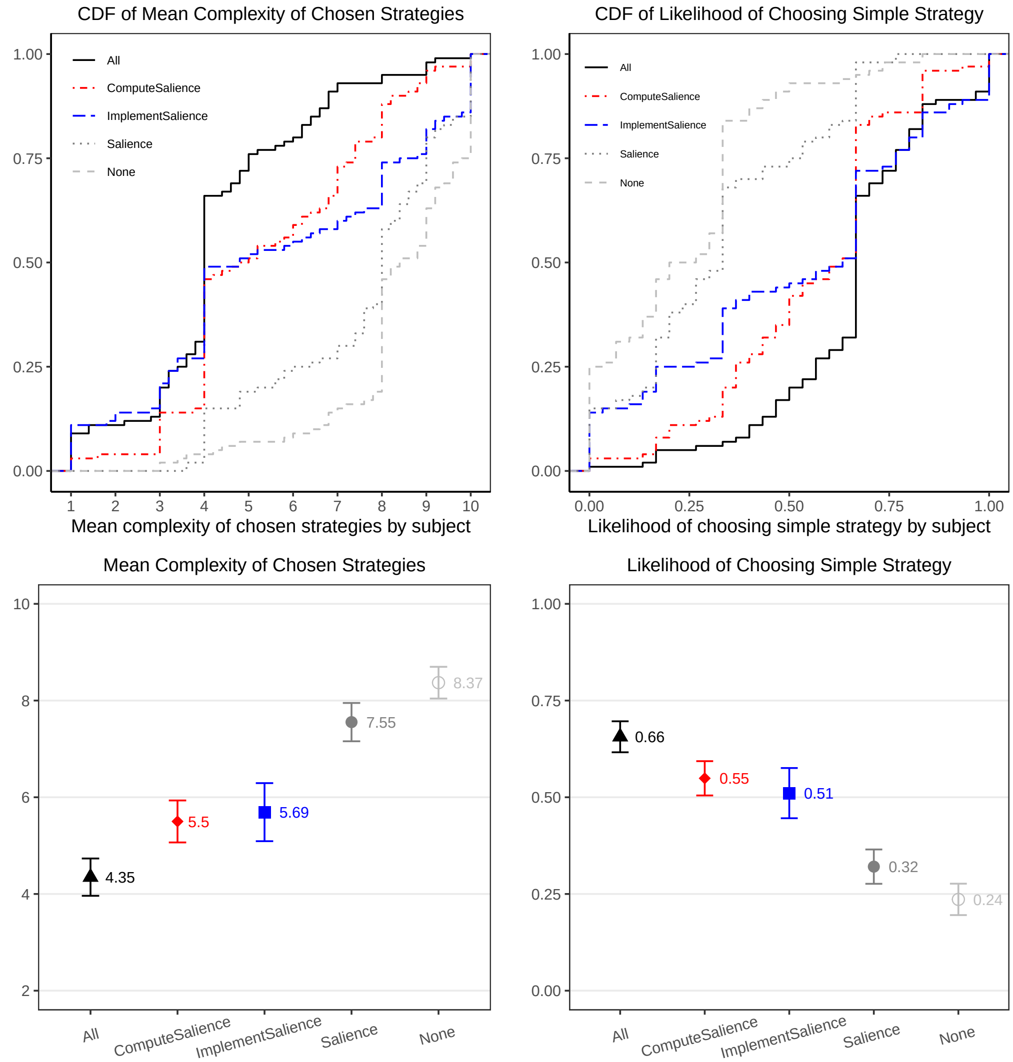

Figure 15: Complexity Level of Chosen Strategies and Likelihood of Choosing Simple Strategy – Last Five Matches Only *Notes: The analysis is on the last five matches. The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. When computing the likelihood of choosing (the relatively) simple strategy in the right panel, a subject's strategy choice in a game is analyzed as six pairwise comparisons. For the bottom two panels, we first compute the statistic for each subject and then calculate the treatment-level averages. The short vertical lines denote 95 percent confidence intervals.*

Table 11: Impacts of Complexity and Strategic Uncertainty on Strategy Choice – Last Five Matches Only

	Strategy Choice (Conditional Logit)		
	(1)	(2)	(3)
complexity	0.341***	0.289***	0.205***
	(0.041)	(0.041)	(0.044)
Salience×complexity	-0.186***	-0.171***	-0.112**
	(0.047)	(0.044)	(0.053)
Computation×complexity	-0.184***	-0.175***	-0.062*
	(0.026)	(0.025)	(0.032)
Implementation×complexity	-0.164***	-0.157***	-0.180***
	(0.026)	(0.024)	(0.031)
BOA		0.562***	1.480***
		(0.202)	(0.383)
Salience×BOA			-0.483
			(0.536)
Computation×BOA			-2.559***
			(0.556)
Implementation×BOA			0.373
			(0.511)
No. of subjects	500	500	500

*Notes: This replicates Table 5 using only the data of the last five matches. Conditional logit regression with the unit of analysis being a strategy choice game as a whole. The dependent variable is an indicator of whether a strategy is chosen. The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. BOA denotes the size of the basin of attraction of a strategy, see Section 2.2. Salience, computation, and implementation are treatment indicators for salience, computational, and implementation complexity, respectively. Clustered standard errors at the subject level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.*

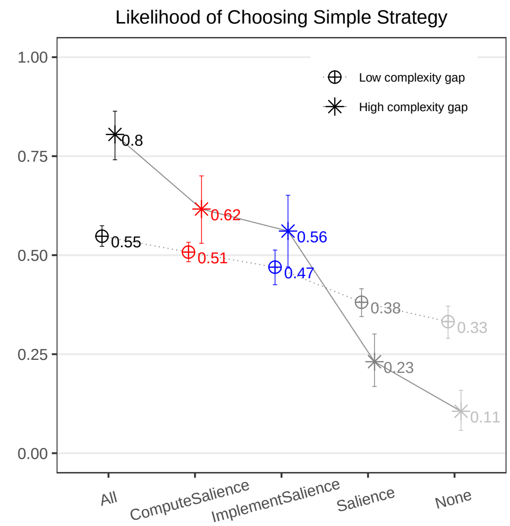

Figure 16: Likelihood of Choosing Simple Strategy by Complexity Gap – Last Five Matches Only *Notes: This replicates Figure 7 using only the data of the last five matches. A subject's strategy choice in a game is analyzed as six pairwise comparisons. Complexity gap is the difference in complexity (i.e., the total number of transitions in the automaton representation) between the two strategies being compared. High (Low) complexity gap means being larger (weakly smaller) than 4, i.e., the median. The short vertical lines denote 95 percent confidence intervals by bootstrapping.*

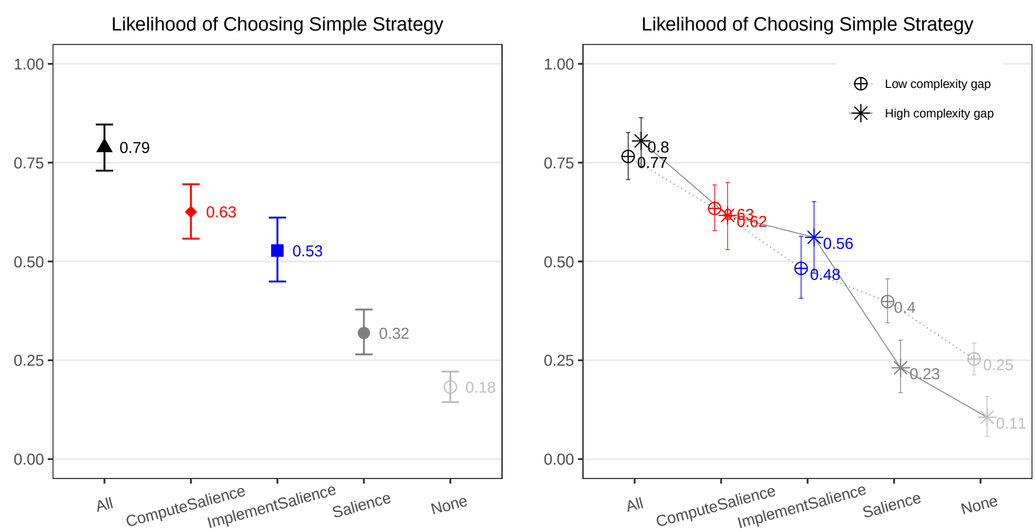

Figure 17: Likelihood of Choosing Simple Strategy – Complex Strategy Having Less Strategic Risk (Last Five Matches Only) *Notes: This replicates Figure 9 using only the data of the last five matches. Analysis is on only the pairwise comparisons in which BOA predicts that the complex strategy is more likely to be chosen. The left panel pools all these pairwise comparisons, while the right panel separately examines pairs with (weakly) lower-than-median complexity gaps (the median is 4) and those with higher-than-median complexity gaps. Complexity gap is the difference in complexity (i.e., the total number of transitions in the automaton representation) between the two strategies being compared. The short vertical lines denote 95 percent confidence intervals by bootstrapping.*

C.2 Laboratory Experiment

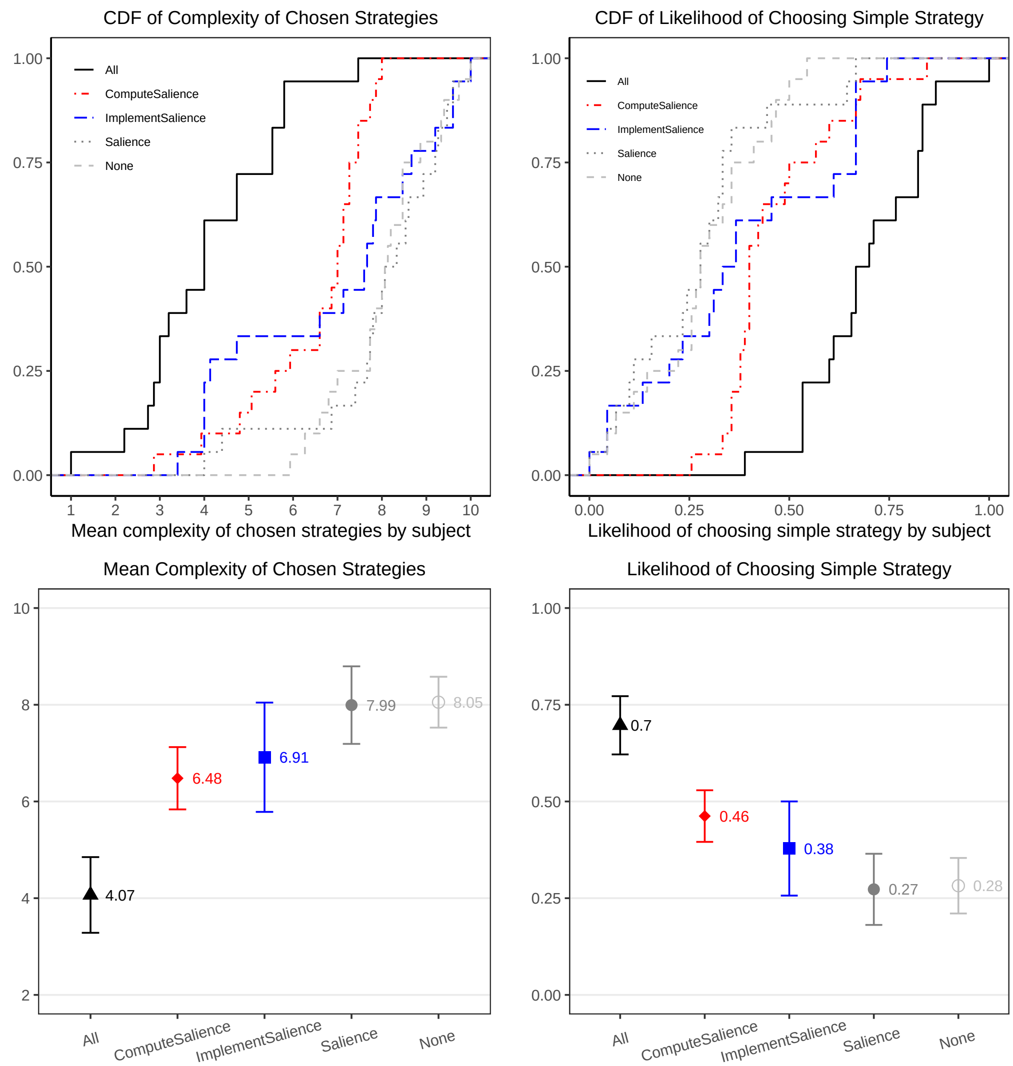

Figure 18: Complexity Level of Chosen Strategies and Likelihood of Choosing Simple Strategy – Laboratory Experiment *Notes: The analysis is on the laboratory experiment in which subjects play the strategy choice game simultaneously. The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. When computing the likelihood of choosing (the relatively) simple strategy in the right panel, a subject's strategy choice in a game is analyzed as six pairwise comparisons. For the bottom two panels, we first compute the statistic for each subject and then calculate the treatment-level averages. The short vertical lines denote 95 percent confidence intervals.*

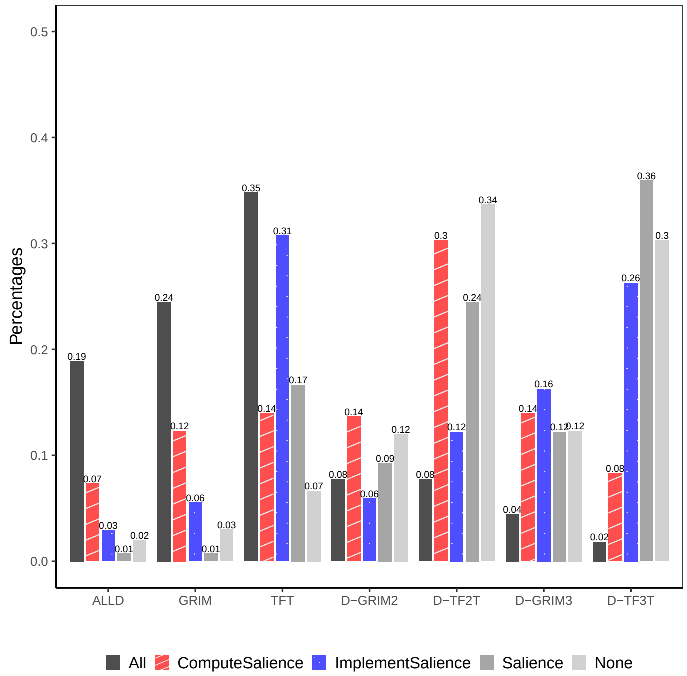

Figure 19: Distribution of Chosen Strategies Across Treatments – Laboratory Experiment
Notes: The analysis is on the laboratory experiment in which subjects play the strategy choice game simultaneously.

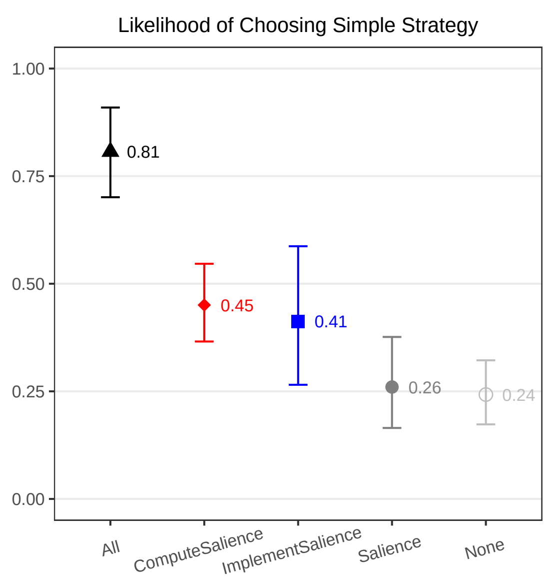

Figure 20: Likelihood of Choosing Simple Strategy – Complex Strategy Having Less Strategic Risk (Laboratory Experiment) *Notes: The analysis is on the laboratory experiment in which subjects play the strategy choice game simultaneously. Analysis is on only the pairwise comparisons in which BOA predicts that the complex strategy is more likely to be chosen. The short vertical lines denote 95 percent confidence intervals by bootstrapping.*

Table 12: Impacts of Complexity and Strategic Uncertainty on Strategy Choice – Laboratory Experiment

	Strategy Choice (Conditional Logit)		
	(1)	(2)	(3)
complexity	0.271***	0.251***	0.221***
	(0.048)	(0.046)	(0.049)
Salience×complexity	0.023	0.024	0.008
	(0.075)	(0.073)	(0.075)
Computation×complexity	-0.261***	-0.255***	-0.152***
	(0.056)	(0.052)	(0.058)
Implementation×complexity	-0.214***	-0.210***	-0.203***
	(0.050)	(0.049)	(0.054)
BOA		0.246	0.618
		(0.355)	(0.723)
Salience×BOA			0.367
			(0.962)
Computation×BOA			-2.065***
			(0.782)
Implementation×BOA			-0.230
			(0.923)
No. of subjects	94	94	94

*Notes: The analysis is on the laboratory experiment in which subjects play the strategy choice game simultaneously. Conditional logit regression with the unit of analysis being a strategy choice game as a whole. The dependent variable is an indicator of whether a strategy is chosen. The complexity of a strategy is measured by the total number of transitions in its automaton representation, see Section 2.1. BOA denotes the size of the basin of attraction of a strategy, see Section 2.2. Salience, computation, and implementation are treatment indicators for salience, computational, and implementation complexity, respectively. Clustered standard errors at the subject level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.*

C.3 Exploitative Strategies

The complex strategies in our strategy choice game are all exploitative strategies. A potential concern is therefore that the aversion to complex strategies we document merely reflects a dislike of exploitative strategies rather than complexity aversion per se. Our data provide two pieces of evidence against this interpretation. First, ALLD, which is an exploitative strategy and the simplest one, is chosen more frequently in All, ComputeSalience, and ImplementSalience than in Salience and then None (see Figure 13 in Appendix B). If subjects were simply averse to exploitation, we would not expect such a systematic pattern. Second, we replicate our main analysis

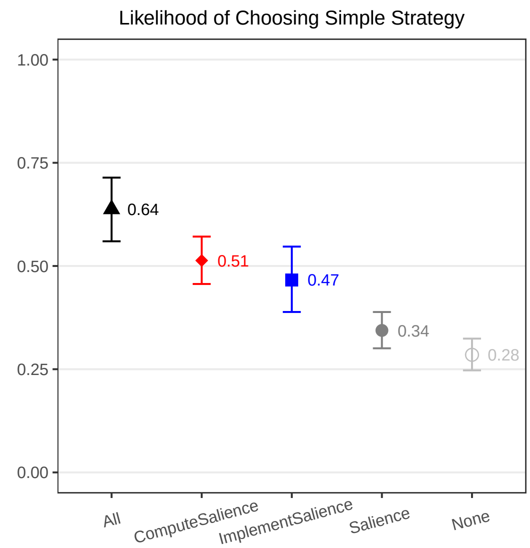

Figure 21: Likelihood of Choosing Simple Strategy – Exploitative Strategies Only *Notes: Analysis is on only the pairwise comparisons among the five exploitative strategies—ALLD, D-GRIM2, D-TF2T, D-GRIM3, and D-TF3T. The short vertical lines denote 95 percent confidence intervals by bootstrapping.*

using only pairwise comparisons among the five exploitative strategies (ALLD, D-GRIM2, D-TF2T, D-GRIM3, and D-TF3T), thereby holding the exploitative nature of strategies constant and isolating variation in complexity alone. As shown in Figure 21, the monotonic pattern of treatment effects is preserved: the likelihood of choosing the relatively simple (exploitative) strategy increasingly rises as more dimensions of complexity are present, changing from 0.28 in None to 0.64 in All. These results confirm that subjects respond to the complexity of strategies, not merely to whether they are exploitative.

D Earlier Experiment

An earlier version of this paper featured a different experiment. We re-designed and reran the experiment in response to a set of excellent referee comments, raising concerns about the interpretability of the original design that we ultimately agreed with. Despite several differences between the designs, the previous experiment generated very similar results and conclusions. To summarize the key differences between the designs:

- The earlier experiment consisted of two sub-experiments: a Non-Strategic experiment in which subjects played against a robot whose strategy was known, and a Strategic experiment in which subjects were matched with human counterparts. The current experiment focuses exclusively on the strategic setting.

- The earlier Strategic experiment had four treatments (analogues of our None, Salience, ImplementSalience, ComputeSalience), while the current experiment has an additional treatment—All, which is closest to the standard repeated prisoner's dilemma experiments, containing all three "faces" of complexity.

- The earlier Strategic experiment had subjects play eleven different strategy choice games that varied in game size (i.e., the number of available strategies, being 2, 3 or 6), strategy menu, continuation probability, and implementation error rate (which was framed as imperfect monitoring).[62] The current experiment uses a single game played 15 times with feedback, allowing for a direct analysis of the robustness of complexity aversion to learning.

- The earlier Strategic experiment studied imperfect monitoring games, in which both the subject's and her counterpart's actions in the Implementation Stage were reversed with a known probability each round and this probability was told to the subject. The current experiment instead studies standard prisoner's dilemmas but injects (a much smaller amount of) realistic noise, estimated from prior behavior in a prisoner's dilemma experiment in order to make measured implementation costs more ecologically valid. In the Implementation Stage, this probabilistic error is injected only into the counterpart's strategy, and subjects are not told what this probability is, but only that it is estimated from other subjects' behavior, requiring them to (as in standard prisoner's dilemmas) form their own beliefs based on their beliefs about the noisiness of other subjects.

- The current experiment adds cognitive test questions, allowing for tests of the robustness of the main findings to cognitive ability. The current experiment also has a larger lab sample, allowing meaningful comparison of lab and Prolific samples.

- The current experiment also features a number of small refinements and adjustments to the instructions and subject feedback that we think make the experiment clearer to subjects.

Despite these design differences, the main findings (using the same analyses as

[62]By varying these dimensions, we constructed eleven games that are either *pure* coordination games with exactly equal equilibrium payoffs or (general coordination) games in which payoffs from different equilibria are very flat.

the current study) from the earlier Strategic experiment are broadly consistent with the current results. Importantly, both studies find that all three facets of complexity influence strategy choice, driving subjects towards the use of algorithmically simpler strategies, and the influence of either computational or implementation complexity dominates that of strategic uncertainty when complexity and strategic uncertainty have opposite predictions on strategy choice. The two minor differences in our current experiment are (i) salience effects are a little stronger than in the earlier experiment and (ii) implementation complexity is weaker relative to computational complexity than in the earlier design.[63]

[63] We think this is likely because the eleven games in our earlier experiment, by design, have either exactly equal or very flat diagonal payoffs, while in the current experiment we have a larger tradeoff between complexity and payoffs (in order to allow us to measure payoff effects as documented in Section 4.3). We suspect that because of this, payoff dominance effects were stronger in our ImplementSalience treatment (where subjects directly see payoffs from strategy combinations) than in the earlier design, reducing the severity of implementation complexity effects.

E Instructions and Screenshots

The following provides the instructions and screenshots of the software interfaces for the Prolific experiment. The lab experiment is essentially identical.

E.1 Instructions

Introduction

- We will start by providing you with **instructions** for the study.
- We will ask you **comprehension questions** to check that you understand the instructions. **If you answer all these questions correctly on your first attempt, you will receive a bonus of $0.5.** Please read and follow the instructions closely and carefully.
- If you complete the study, you will receive a **guaranteed payment** of $7.00. In addition, **20% of all the participants** will be randomly selected and receive also a sizable performance-based bonus payment that is based on their choices in the **Decision** portion of the study.
- You will play **15 games in the main part of the study**. In each game, you will be **randomly matched with another participant**. If you are selected to receive also a performance-based bonus payment, then at the end of the study, one game will be **randomly picked** (each game is equally likely to be picked), and you will **get $0.5 for every point** you earn in that picked game.

The Game: Descriptions

- Each game is divided up into a number of **rounds**. In each round, you will choose an action, either '**A**' or '**B**.' Your counterpart will also choose an action, either '**A**' or '**B**'. The **combination** of your choice and your counterpart's choice each round will determine **how many points** you earn for that round. We will show you these payoffs in a **Payment Table** like the one below:

Payment Table

		Counterpart's Action	
		A	B
Your Action	A	0, 0	8, -2
	B	-2, 8	6, 6

- Your action choice will determine the row of the table, your counterpart's choice will determine the column. The first number (in purple) shows your resulting payoff and the second number (in orange) shows your counterpart's. (Your counterpart will see the same information when making their choice.)
- In the example above, when you and your counterpart both choose **A** each of you earns 0 point for the round, but when you both choose **B** each of you earns 6 points. When you choose **A** and your counterpart chooses **B** you earn 8 points and your counterpart loses 2 points for the round, but in the reverse case you lose 2 points and your counterpart earns 8 points.
- The **number of rounds** in each game (the "game length") is **randomly determined** by the computer in a simple way: each round there is a 1/8 chance that the computer will end the game and a 7/8 chance that the game will continue for additional rounds.

Comprehension Questions

The number of rounds in each game:
○ is equal to 1.
○ is equal to 8.
○ is determined randomly, with the game continuing with a 7/8 chance after each round.

The *All* and *ComputeSalience* treatments.

Choose Your Rule

- Instead of directly choosing your actions (A or B) during each round of the game, you will **choose a rule ahead of time** to determine what actions you will take in response to your counterpart's actions in the game.
- Your counterpart has also chosen a rule to determine what actions he/she would like to take in response to your actions in the game. (Technically, we collected your counterpart's rule choices in an earlier experiment, so your counterpart made choices to respond to his/her counterpart's choices.)
- You, however, will not be told your counterpart's rule when choosing your rule (and he/she was not told his/her counterpart's rule either).
- Before each game begins you will see a **Rules Table** as shown in the below example. The Rules Table will show you:
  1. **what rules your counterpart can choose from** in that game, in the columns (shown in orange)
  2. **the same set of rules you can choose from**, in the rows (shown in purple).

Example

The Payment Table

Counterpart's Action

		A	B
Your Action	A	0, 0	8, -2
	B	-2, 8	6, 6

The Rules Table

(Mouse over any rule to see it in a larger font.)

Counterpart's Rule

	Rule J: Always choose A.	Rule H: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B.	Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.	Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.	Rule T: Choose B until you see your counterpart chose A, after which switch to A forever.	Rule D: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever.	Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.
○ Rule J: Always choose A.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
○ Rule H: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
○ Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Your Rule ○ Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.	?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
○ Rule D: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
○ Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?

- When looking at the table, you can mouse over the rules to see them better. To choose a rule, just click the button next to the rule. You will not get to see what rule your counterpart chose when making your own rule choice (the same was true for your counterpart).
- Your **base payment** in a game: After you have both chosen your rules, the computer will pay you the amount you would earn **on average** if you and your counterpart followed the chosen rules in **one million of games** (each game has a random length), given the numbers in the Payment Table. In order to make this more realistic, when making these calculations we will include a small chance (based on estimates from previous participants' behavior in games in which they directly chose their actions each round) that you and your counterpart accidentally choose the wrong action each round. Thus, the computer will pay you the **average amount** you can expect to earn in a game, **given the rules** each of you has chosen, the numbers in the Payment Table and realistic estimates of the likelihood you would make errors when implementing these rules. (Technically, your counterpart was paid in the exact same way you will be, based on the rule he/she chose in response to his/her beliefs about his/her counterpart's unknown rule choice.)
- This may sound complicated, but what it means for you is actually relatively simple. **When selecting a rule, you should think about how much following the rule would earn you 'on average'** (over all the possible random game lengths) given the numbers in the Payment Table and **the rule you think your counterpart chose when playing the same game**.
- Inside the **Rules Table** we show you question marks (?), for every combination of (i) your rule choice (rows, in purple) and (ii) your counterpart's rule (column, in orange). These are to remind you that (i) each combination of your and your counterpart's rules pay a determinate amount (the average over many possible games) but (ii) we will not show this amount to you - this is for you to figure out. The same was true for your counterpart.

Comprehension Question

Your base payment in each game is determined by:

○ your and your counterpart's direct choices of A versus B in each of a series of rounds.

○ your and your counterpart's chosen rules. Specifically, the computer will pay you the amount you would earn if you and your counterpart followed the chosen rules, in a single game.

○ your and your counterpart's chosen rules. Specifically, the computer will pay you the amount you would earn if you and your counterpart followed the chosen rules, averaged over one million of games with randomly different lengths.

The *None* treatment.

Choose Your Rule

- Instead of directly choosing your actions (A or B) during each round of the game, you will **choose a rule ahead of time** to determine what actions you will take in response to your counterpart's actions in the game. An example of a rule is "Always choose A" or "Choose B in round 1 and then choose whatever your counterpart chose in the previous round".
- Your counterpart has also chosen a rule to determine what actions he/she would like to take in response to your actions in the game. (Technically, we collected your counterpart's rule choices in an earlier experiment, so your counterpart made choices to respond to his/her counterpart's choices.)
- You, however, will not be told your counterpart's rule when choosing your rule (and he/she was not told his/her counterpart's rule either).
- We are not going to tell you the details of the rules you can choose from (but each rule determines what actions you will take in response to your counterpart's actions in the game); instead, we will give each of them a code (Rule Q, Rule J, etc.). The same was true for your counterpart.
- Before each game begins you will see a **Rules Table** as shown in the below example. The Rules Table will show you:
  1. **what rules your counterpart can choose from** in that game, in the columns (shown in orange)
  2. **the same set of rules you can choose from**, in the rows (shown in purple).

Example

The Payment Table

		Counterpart's Action A	Counterpart's Action B
Your Action	A	0, 0	8, -2
	B	-2, 8	6, 6

The Rules Table

		Counterpart's Rule: Rule Q	Rule J	Rule D	Rule K	Rule H	Rule L	Rule T
Your Rule	○ Rule Q	40, 40	0, 13	38, 39	38, 39	40, 40	35, 37	7, 16
	○ Rule J	13, 0	2, 2	15, -1	12, 0	20, -2	9, 1	10, 1
	○ Rule D	39, 38	-1, 15	40, 40	36, 35	40, 40	35, 36	8, 14
	○ Rule K	39, 38	0, 12	35, 36	37, 37	40, 39	8, 17	27, 30
	○ Rule H	40, 40	-2, 20	40, 40	39, 40	40, 40	35, 38	6, 22
	○ Rule L	37, 35	1, 9	36, 35	17, 8	38, 35	35, 35	9, 8
	○ Rule T	16, 7	1, 10	14, 8	30, 27	22, 6	8, 9	28, 28

- To choose a rule, just click the button next to the rule. You will not get to see what rule your counterpart chose when making your own rule choice (the same was true for your counterpart).
- Your **base payment** in a game: After you have both chosen your rules, the computer will pay you the amount you would earn **on average** if you and your counterpart followed the chosen rules in **one million of games** (each game has a random length), given the numbers in the Payment Table. In order to make this more realistic, when making these calculations we will include a small chance (based on estimates from previous participants' behavior in games in which they directly chose their actions each round) that you and your counterpart may accidentally choose the wrong action each round. Thus, the computer will pay you the **average amount** you can expect to earn in a game, **given the rules** each of you has chosen, the numbers in the Payment Table and realistic estimates of the likelihood you would make errors when implementing these rules. (Technically, your counterpart was paid in the exact same way as you will be, based on the rule he/she chose in response to his/her beliefs about his/her counterpart's unknown rule choice.)
- This may sound complicated, but what it means for you is actually relatively simple. **When selecting a rule, you should think about how much following the rule would earn you 'on average'** (over all the possible random game lengths) given the numbers in the Payment Table and **the rule you think your counterpart chose when playing the same game**.
- Inside the **Rules Table** we show you what exactly these average amounts (i.e., the base payments) are, for every combination of (i) your rule choice (rows, in purple) and (ii) your counterpart's rule (column, in orange).

  Comprehension Question

  Your base payment in each game is determined by:

  ○ your and your counterpart's direct choices of A versus B in each of a series of rounds.

  ○ your and your counterpart's chosen rules. Specifically, the computer will pay you the amount you would earn if you and your counterpart followed the chosen rules, in a single game.

  ○ your and your counterpart's chosen rules. Specifically, the computer will pay you the amount you would earn if you and your counterpart followed the chosen rules, averaged over one million of games with randomly different lengths.

The *ImplementSalience* and *Salience* treatments.

Choose Your Rule

- Instead of directly choosing your actions (A or B) during each round of the game, you will **choose a rule ahead of time** to determine what actions you will take in response to your counterpart's actions in the game.
- Your counterpart has also chosen a rule to determine what actions he/she would like to take in response to your actions in the game. (Technically, we collected your counterpart's rule choices in an earlier experiment, so your counterpart made choices to respond to his/her counterpart's choices.)
- You, however, will not be told your counterpart's rule when choosing your rule (and he/she was not told his/her counterpart's rule either).
- Before each game begins you will see a **Rules Table** as shown in the below example. The Rules Table will show you:
  1. **what rules your counterpart can choose from** in that game, in the columns (shown in orange)
  2. **the same set of rules you can choose from**, in the rows (shown in purple).

Example

The Payment Table

		Counterpart's Action	
		A	B
Your Action	A	0, 0	8, -2
	B	-2, 8	6, 6

The Rules Table

(Mouse over any rule to see it in a larger font.)

Your Rule	Counterpart's Rule: Rule D: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever.	Rule H: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B.	Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.	Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.	Rule J: Always choose A.	Rule T: Choose B until you see your counterpart chose A, after which switch to A forever.	Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.
○ Rule D: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever.	40, 40	40, 40	36, 35	39, 38	-1, 15	8, 14	35, 36
○ Rule H: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B.	40, 40	40, 40	39, 40	40, 40	-2, 20	6, 22	35, 38
○ Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.	35, 36	40, 39	37, 37	39, 38	0, 12	27, 30	8, 17
○ Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.	38, 39	40, 40	38, 39	40, 40	0, 13	7, 16	35, 37
○ Rule J: Always choose A.	15, -1	20, -2	12, 0	13, 0	2, 2	10, 1	9, 1
○ Rule T: Choose B until you see your counterpart chose A, after which switch to A forever.	14, 8	22, 6	30, 27	16, 7	1, 10	28, 28	8, 9
○ Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.	36, 35	38, 35	17, 8	37, 35	1, 9	9, 8	35, 35

- When looking at the table, you can mouse over the rules to see them better. To choose a rule, just click the button next to the rule. You will not get to see what rule your counterpart chose when making your own rule choice (the same was true for your counterpart).
- Your **base payment** in a game: After you have both chosen your rules, the computer will pay you the amount you would earn **on average** if you and your counterpart followed the chosen rules in **one million of games** (each game has a random length), given the numbers in the Payment Table. In order to make this more realistic, when making these calculations we will include a small chance (based on estimates from previous participants' behavior in games in which they directly chose their actions each round) that you and your counterpart accidentally choose the wrong action each round. Thus, the computer will pay you the **average amount** you can expect to earn in a game, **given the rules** each of you has chosen, the numbers in the Payment Table and realistic estimates of the likelihood you would make errors when implementing these rules. (Technically, your counterpart was paid in the exact same way you will be, based on the rule he/she chose in response to his/her beliefs about his/her counterpart's unknown rule choice.)
- This may sound complicated, but what it means for you is actually relatively simple. **When selecting a rule, you should think about how much following the rule would earn you 'on average'** (over all the possible random game lengths) given the numbers in the Payment Table and **the rule you think your counterpart chose when playing the same game**.
- Inside the **Rules Table** we show you what exactly these average amounts (i.e., the base payments) are, for every combination of (i) your rule choice (rows, in purple) and (ii) your counterpart's rule (column, in orange).

Comprehension Question

Your base payment in each game is determined by:

○ your and your counterpart's direct choices of A versus B in each of a series of rounds.

○ your and your counterpart's chosen rules. Specifically, the computer will pay you the amount you would earn if you and your counterpart followed the chosen rules, in a single game.

○ your and your counterpart's chosen rules. Specifically, the computer will pay you the amount you would earn if you and your counterpart followed the chosen rules, averaged over one million of games with randomly different lengths.

This page is shown in all treatments except *None*.

The Rules You Can Choose From

In each game, the set of rules you and your counterpart can choose from are **always the same set of 7 rules** , each of which specifies **which action you (your counterpart) will choose** in response to **your counterpart's (your) actions** in past rounds of the game. We will describe the 7 rules and explain how they work in an example below.

The 7 Rules and How They Work

Example: Suppose your counterpart's actions are AAAB in the first 4 rounds of a game (i.e., he/she chose A, then A, then A, then B). Here are the actions you would take in response if you had chosen each of the 7 rules.

(1) If you choose **"Rule J: Always choose A."**

- then you will respond in rounds 1-5 with actions AAAAA. That is, you always choose A.

(2) If you choose **"Rule H: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B."**

- then you will respond in rounds 1-5 with actions ABBAB. That is, you choose A in round 1; in rounds 2-3, you choose B; in round 4, you choose A since your counterpart chose AAA in the previous 3 rounds (i.e., rounds 1-3); in round 5, you choose B since your counterpart didn't choose AAA in the previous 3 rounds (i.e., rounds 2-4).

(3) If you choose **"Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever."**

- then you will respond in rounds 1-5 with actions ABAAA. That is, you choose A in round 1; in round 2, you choose B; in round 3, you switch to A since your counterpart chose AA in the previous 2 rounds (i.e., rounds 1-2); you then keep choosing A forever.

(4) If you choose **"Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B."**

- then you will respond in rounds 1-5 with actions ABAAB. That is, you choose A in round 1; in round 2, you choose B; in round 3 (round 4), you choose A since your counterpart chose AA in the previous 2 rounds — rounds 1-2 (rounds 2-3); in round 5, you choose B since your counterpart didn't choose AA in the previous 2 rounds (i.e., rounds 3-4).

(5) If you choose **"Rule T: Choose B until you see your counterpart chose A, after which switch to A forever."**

- then you will respond in rounds 1-5 with actions BAAAA. That is, you choose B in round 1; in round 2, you switch to A since your counterpart chose A in round 1; you then keep choosing A forever.

(6) If you choose **"Rule D: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever."**

- then you will respond in rounds 1-5 with actions ABBAA. That is, you choose A in round 1; in rounds 2-3, you choose B; in round 4, you switch to A since your counterpart chose AAA in the previous 3 rounds (i.e., rounds 1-3); you then keep choosing A forever.

(7) If you choose **"Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round."**

- then you will respond in rounds 1-5 with actions BAAAB. That is, you choose B in round 1; in round 2, you choose what your counterpart chose in round 1, A; in round 3, you choose what your counterpart chose in round 2, A; then go on.

Subjects need to click buttons to see the rules one by one.

This page is shown in *All* and *ImplementSalience* treatments.

Following Your Rule

- As just discussed, your **base payment** for each game will be the **average amount** your rule and your counterpart's rule would earn you **over one million of games of varying lengths**. However, we will also have you do a **second task** in each game that also depends on the rule you choose for the game. The same was true for your counterpart.

- In each game, after you have chosen your rule, we will have you actually **follow your rule** over a series of rounds, on a screen like the one below. That is, you will have to make the action choices that your chosen rule specifies **in response to the actions specified by your counterpart's chosen rule**, which we refer to as your "Counterpart's Action". For your counterpart's actions, we will again include a small chance (based on estimates of previous participants' behavior) each round that your counterpart's rule might be wrongly implemented.

- Each round, your counterpart's actions (in orange) and your actions (in purple) in past rounds will always be shown to you on the screen.

Example

Your Rule:

Your chosen rule in the game will be shown here.

Your Action Choice in this Round (Round #)

Following your rule, which action should you choose:

A B

Round History

Round	Counterpart's Action	Your Action
...	...	...
3	B	A
2	A	B
1	B	A

- Your job is to **accurately follow the rule you've chosen**: we will penalize you for each mistake you make in following that rule. **Every time you make an action choice that conflicts what your chosen rule instructs you to do, you will lose 1 point** from your base payment in the game. In other words, your earnings from the game will be your base payment minus your losses due to mistakes you make in following your chosen rule.

- REMEMBER: your choice of a rule has two effects on your payments. First, it influences your **base payment** in the game as discussed before. Second (and separately), your choice of rule determines what rule you have to actually follow in the second part of the game. After choosing your rule and determining your base payment, the only impact your actions have on your earnings for the rest of the game is to produce possible **penalties** to your base payment **if you make mistakes** in following the rule.

Comprehension Question

Your earnings in each game:

○ will be your base payment (determined by your and your counterpart's chosen rules).
○ will be your base payment (determined by your and your counterpart's chosen rules) minus 1 point for each mistake you made when following your chosen rule.
○ will be based entirely on your and your counterpart's actions (A or B) each round in the stage of following rules.

The *All* treatment.

Summary

- You will choose a rule for playing the game shown in the **Payment Table** against your counterpart (who also chose a rule for the same game) for a random number of rounds (each game continues with a 7/8 chance after each round). You both make this choice without knowing what rules your counterparts chose.
- The combination of your and your counterpart's rules will be used (by the computer) to calculate how much you would earn, **on average**, over one million of games (each has a random number of rounds). The average amount will be your **base payment** in the game.
- The **Rules Table** shows you question marks. These are to remind you that (i) each combination of your and your counterpart's rules pay a determinate amount (i.e., base payment) but (ii) we will not show this amount to you - this is for you to figure out. The same was true for your counterpart.
- After choosing your rule, you will have to correctly follow the rule. You will **lose 1 point from your base payment for every mistake** you make in following your rule. In other words, your earnings from the game will be your base payment minus your loss due to mistakes you make.
- You will play **15 games in the main part of the study**. In each game, you will be **randomly matched with another participant**. If you are selected to receive also a performance-based bonus payment, at the end of the study, one game will be **randomly picked** (each game is equally likely to be picked), and you will **get $0.5 for every point** you earn in that picked game.
- Click "**Continue**" to start Game 1.

The *None* and *Salience* treatments.

Summary

- You will choose a rule for playing the game shown in the **Payment Table** against your counterpart (who also chose a rule for the same game) for a random number of rounds (each game continues with a 7/8 chance after each round). You both make this choice without knowing what rules your counterparts chose.
- The combination of your and your counterpart's rules will be used (by the computer) to calculate how much you would earn, **on average**, over one million of games (each has a random number of rounds). The average amount will be your **base payment** in the game.
- The **Rules Table** will show you exactly how much each combination of your and your counterpart's rules will earn you. (Your counterpart saw the same information when choosing his/her rule.)
- After choosing your rule, **the computer will automatically perform the rule you have chosen for you** against your counterpart's rule. And you earn your **base payment** in the game.
- You will play **15 games in the main part of the study**. In each game, you will be **randomly matched with another participant**. If you are selected to receive also a performance-based bonus payment, at the end of the study, one game will be **randomly picked** (each game is equally likely to be picked), and you will **get $0.5 for every point** you earn in that picked game.
- Click "**Continue**" to start Game 1.

The versions for *ComputeSalience* and *ImplementSalience* are variations (specifically in terms of the third and fourth bullet points) of the two shown versions.

E.2 Screenshots

The *All* treatment.

Game 1– Choose Your Rule

The Payment Table

Your Action \ Counterpart's Action	A	B
A	0, 0	8, -2
B	-2, 8	6, 6

The Rules Table

(Mouse over any rule to see it in a larger font.)

Counterpart's Rule

Your Rule	Rule J: Always choose A.	Rule H: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B.	Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.	Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.	Rule T: Choose B until you see your counterpart chose A, after which switch to A forever.	Rule D: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever.	Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.
Rule J: Always choose A.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Rule H: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 3 rounds; otherwise choose B.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Rule Q: In round 1, choose A. After round 1: choose A if your counterpart chose A in each of the previous 2 rounds; otherwise choose B.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Rule T: Choose B until you see your counterpart chose A, after which switch to A forever.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Rule D: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 3 rounds, after which switch to A forever.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?
Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?	?, ?

Please choose your rule and then click Submit.

Submit

The *None* treatment.

Game 1– Choose Your Rule

The Payment Table

Counterpart's Action

Your Action	A	B
A	0, 0	8, -2
B	-2, 8	6, 6

The Rules Table

Counterpart's Rule

Your Rule	Rule Q	Rule J	Rule D	Rule K	Rule H	Rule L	Rule T
○ Rule Q	40, 40	0, 13	38, 39	38, 39	40, 40	35, 37	7, 16
○ Rule J	13, 0	2, 2	15, -1	12, 0	20, -2	9, 1	10, 1
○ Rule D	39, 38	-1, 15	40, 40	36, 35	40, 40	35, 36	8, 14
○ Rule K	39, 38	0, 12	35, 36	37, 37	40, 39	8, 17	27, 30
○ Rule H	40, 40	-2, 20	40, 40	39, 40	40, 40	35, 38	6, 22
○ Rule L	37, 35	1, 9	36, 35	17, 8	38, 35	35, 35	9, 8
○ Rule T	16, 7	1, 10	14, 8	30, 27	22, 6	8, 9	28, 28

Please choose your rule and then click Submit.

Submit

(REMINDER: The computer will automatically perform your chosen rule for you against your counterpart's rule.)

The interfaces for the other three treatments are variations between the interfaces for *All* and *None* treatments.

In treatments with implementation complexity:

Game 1, Round 3 – Choose Your Action!

Your Rule:

Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.

Your Action Choice in this Round (Round 3)

Following your rule, which action should you choose? (Reminder: At this stage, your job is only to correctly follow the rule you have chosen.)

A B

Round History

Round	Counterpart's Action	Your Action
3	?	?
2	B	A
1	A	B

When making a mistake in implementation:

Game 1, Round 3 – You Made a Mistake!

Your Rule:

Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.

Mistake in this Round (Round 3)

Your chosen action in this round didn't follow your chosen rule! You lost 1 point!

Next

Round History

Round	Counterpart's Action	Your Action
3	B	A
2	B	A
1	A	B

In treatments with implementation complexity, at the end of a game:

Game 1 – The current game has ended!

Your Rule:

Rule K: In round 1, choose B. After round 1: choose whatever your counterpart chose in the previous round.

You made **1** mistake(s) in implementing your rule and thus lost 1 point from your base payment (which is determined by your and your counterpart's chosen rules).
So your payoff from this game is 35 points for the base payment minus 1 point for mistake(s), which equals **34 points**.

Please click the "Next" button to start a new game. **You will be randomly matched with another participant in the new game.**

Next

History of Rule Implementation

Round	Counterpart's Action	Your Action
6	B	B
5	B	B
4	B	B
3	B	A
2	B	A
1	A	B

Game 2 – The current game has ended!

Your Rule:

Rule L: In round 1, choose A. After round 1: choose B until you see your counterpart chose A in each of the previous 2 rounds, after which switch to A forever.

You made **0** mistake in implementing your rule. So your payoff from this game is exactly your base payment (which is determined by your and your counterpart's chosen rules), **17 points**.

Please click the "Next" button to start a new game. **You will be randomly matched with another participant in the new game.**

Next

History of Rule Implementation

Round	Counterpart's Action	Your Action
4	B	B
3	B	B
2	A	B
1	B	A

In treatments without implementation complexity:

Game 1 – The current game has ended!

Your Rule:

Rule J: Always choose A.

You earn your base payment (which is determined by your and your counterpart's chosen rules) in the game, **13 points**.

Please click the "Next" button to start a new game. **You will be randomly matched with another participant in the new game.**

Next

Note that in the *None* treatment, subjects see only the rule label.