

Zhu, Jian-Qiao; Peterson, Joshua C.; Enke, Benjamin; Griffiths, Thomas L.

**Working Paper**

## Capturing the Complexity of Human Strategic Decision-Making with Machine Learning

CESifo Working Paper, No. 11296

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Zhu, Jian-Qiao; Peterson, Joshua C.; Enke, Benjamin; Griffiths, Thomas L. (2024) : Capturing the Complexity of Human Strategic Decision-Making with Machine Learning, CESifo Working Paper, No. 11296, CESifo GmbH, Munich

This Version is available at:

<https://hdl.handle.net/10419/305538>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Capturing the Complexity of Human Strategic Decision- Making with Machine Learning

*Jian-Qiao Zhu, Joshua C. Peterson, Benjamin Enke, Thomas L. Griffiths*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: <https://www.cesifo.org/en/wp>

# Capturing the Complexity of Human Strategic Decision-Making with Machine Learning

## Abstract

Understanding how people behave in strategic settings—where they make decisions based on their expectations about the behavior of others—is a longstanding problem in the behavioral sciences. We conduct the largest study to date of strategic decision-making in the context of initial play in two-player matrix games, analyzing over 90,000 human decisions across more than 2,400 procedurally generated games that span a much wider space than previous datasets. We show that a deep neural network trained on these data predicts people’s choices better than leading theories of strategic behavior, indicating that there is systematic variation that is not explained by those theories. We then modify the network to produce a new, interpretable behavioural model, revealing what the original network learned about people: their ability to optimally respond and their capacity to reason about others are dependent on the complexity of individual games. This context-dependence is critical in explaining deviations from the rational Nash equilibrium, response times, and uncertainty in strategic decisions. More broadly, our results demonstrate how machine learning can be applied beyond prediction to further help generate novel explanations of complex human behavior.

Keywords: behavioural game theory, large scale experiment, machine learning, behavioral economics, complexity.

*Jian-Qiao Zhu*

*Department of Computer Science  
Princeton University / NJ / USA  
jz5204@princeton.edu*

*Joshua C. Peterson*

*Faculty of Computing and Data Science  
Boston University / MA / USA  
peterson.c.joshua@gmail.com*

*Benjamin Enke*

*Department of Economics  
Harvard University, Cambridge / MA / USA  
enke@fas.harvard.edu*

*Thomas L. Griffiths*

*Department of Computer Science &  
Department of Psychology  
Princeton University / NJ / USA  
tomg@princeton.edu*

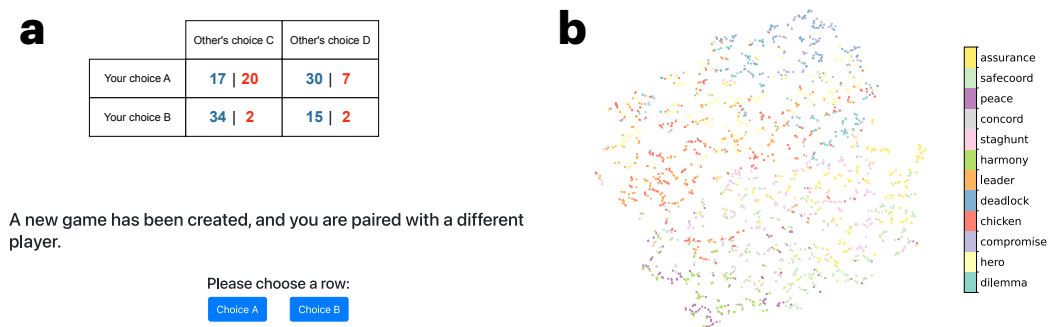
Strategic decision-making is essential when people's outcomes depend on both their own and other people's actions. As a consequence, it is an important topic in various disciplines within the social sciences, including economics, psychology, political science, and artificial intelligence, as well as cultural and biological evolution (Baker et al., 2017; Camerer, 2011; Fehr & Schmidt, 1999; Rawls, 1971; Von Neuman & Morgenstern, 1944; Wright & Leyton-Brown, 2017). The most widely studied type of game in the social sci-

ences is the class of  $2 \times 2$  matrix games (see Figure 1a), which have been found to illuminate behavior in contexts including, among many others, human cooperation and the evolution of morality (Bloom, 2012), price setting and production decisions by firms (Fudenberg & Tirole, 1991), the coordination of investment decisions (Bell & Cover, 1988), and the positioning of political candidates (Morrow, 1994).

The rational model of strategic decisions – the Nash equilibrium – is based on two key assumptions: mutual consistency in beliefs about opponents’ strategies and mutual rationality in best responding to those beliefs. However, despite the prevalent use of Nash equilibria in analyzing matrix games, research has shown that human players often violate both of these assumptions (Camerer, 2011; Gächter, 2004; McKelvey & Palfrey, 1995; Weizsäcker, 2003). Consequently, the effectiveness of these equilibria in explaining people’s strategic behaviors is limited (McKelvey & Palfrey, 1992, 1995). This has prompted the development of behavioral game theory, which has identified various extensions and refinements that produce a closer match to human decisions (Eyster, 2019; Fehr & Schmidt, 1999; Fudenberg & Liang, 2019; Golman et al., 2020).

Despite a proliferation of behavioral models, evaluating the performance of these models has relied on relatively small datasets based on a select group of games, even when combined across different datasets and papers (Fudenberg & Liang, 2019; Wright & Leyton-Brown, 2017). As a result, it remains unclear how well the most popular models of strategic decision-making perform in general. For instance, even seemingly “simple” types of strategic interaction can differ widely in the cognitive difficulty they pose for the actors, yet our understanding of how game complexity shapes behavior is limited. To explore these questions, we conducted a large-scale study of strategic behavior by densely sampling the enormous space of  $2 \times 2$  game structures. We use the resulting dataset to assess the explanatory power of leading behavioral models, quantifying their prediction performance against a machine learning model trained on the same data. This strategy allows us to identify systematic variation that is not captured by existing models, leading us to develop a new interpretable model that captures human behavior almost as well as the machine learning algorithm.

We procedurally generated a dataset of 2,416 matrix games involving monetary gains (see Figure 1b), significantly expanding the diversity of game scenarios studied in prior datasets (a 17-fold increase in the number of games tested relative to the largest meta-analysis in the literature (Wright & Leyton-Brown, 2017)). To systematically sample game matrices, our game generation algorithm is based on the Robinson and Goforth topology for  $2 \times 2$  games, which is constructed using ordinal order graphs of payoffs (Robinson & Goforth, 2005). Each player has 12 unique order graphs in their respective payoff matrices, representing the different ways to rank payoffs between their two strategies. Consequently, there are  $12 \times 12 = 144$  possible game types, considering each player’s order graph independently. We populated all types with at least one pure-strategy Nash equilibrium with procedurally-generated game matrices (see Supplementary Information for details). To investigate human behavior in these settings, we recruited 4,900 participants via Prolific, each of whom was instructed to participate in 20 distinct games, sampled randomly with-



**Figure 1**

**Matrix games.** (a) An example game interface presented to participants, who acted as the row player in each  $2 \times 2$  game. The blue numbers represent the payoffs for the row player, who chooses between strategies A and B, while the red numbers represent the payoffs for the column player, who chooses between strategies C and D. (b) Visualization of game space. Each game is uniquely represented by an 8-integer vector, corresponding to the payoffs to the two players under different configurations of choices. We used *t*-distributed stochastic neighbor embeddings (Van der Maaten & Hinton, 2008) to visualize the spatial relationship between games in a 2D plot, using the Euclidean distance between the embeddings of our best-performing neural network model. Points represent individual games. The colors represent the game topology specific to the row player following Robinson and Goforth (2005).

out replacement from the pool of procedurally-generated games. In total, 93,460 strategic decisions were recorded. No feedback was provided to participants between games, and the players were randomly rematched after each game. Thus, the observed behaviors can be interpreted as initial game play strategies.

We used the resulting dataset to evaluate various models of strategic decision-making with a train-validation split. We fit all models by minimizing the mean squared error (MSE) between model predictions and empirical choice frequencies at the game level on the training set, and then assess their performance in a validation set (see Supplementary Information for details).

The models are based on three key insights from the behavioral game theory literature, each representing a well-established aspect of human cognition in strategic interactions. First, the players may have limited strategic sophistication and compute the best-response function only a small number of times, as captured in level- $k$  models and similar approaches (Camerer et al., 2004; Crawford et al., 2013). Second, a player may exhibit noise in their decision making and also take into account the noisiness of others' behavior, as in quantal-response equilibrium models (McKelvey & Palfrey, 1995). Third, players may be risk averse, meaning they prefer less uncertainty (Fudenberg & Liang, 2019; Murnighan et al., 1988).

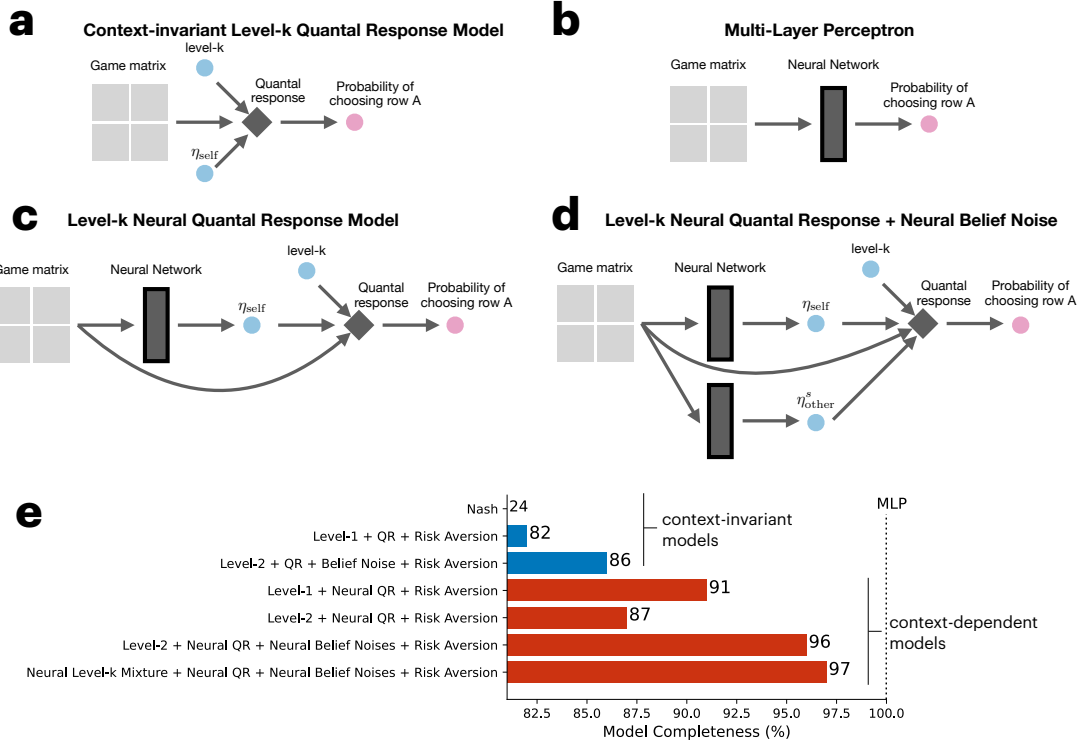


Figure 2

**Model comparisons.** (a) The context-invariant level-k quantal-response model involves three parameters that do not vary across games: strategic sophistication (i.e.,  $k$ ), the players' noisiness (i.e.,  $\eta_{self}$ ) and risk aversion. (b) A Multi-layer Perceptron (MLP) model directly uses the game matrix as input to estimate choice probabilities, without imposing any specific game-theoretic decision-making structure. (c) The level-k neural quantal-response model is a context-dependent model allowing the  $\eta_{self}$  parameter to vary across games. It employs an MLP model, which uses the game matrix as input to estimate game-specific  $\eta_{self}$ . (d) The level-k neural quantal-response and neural belief noise model extends the model in (c) by further learning the game-specific  $\eta_{other}^s$  through two MLP models, each of which takes the game matrix as input. (e) Context-dependent models, incorporating at least one neural network component that allows some or all model parameters to vary across games, outperform context-invariant models in terms of completeness. Higher completeness indicates greater predictive accuracy for human behaviors, with 100% completeness matching the predictive accuracy of the MLP model. All reported results were based on 10-fold cross-validation (see Supplementary Information for details). Our focus was on the heterogeneity across games rather than the heterogeneity across participants.

All models that we estimate are variants of our baseline behavioral model, which is a risk-averse level-k quantal response (QR) model (see Figure 2a) (Wright & Leyton-



Brown, 2017). According to this model, a player forms beliefs about their opponent's behavior (albeit possibly imperfectly) and selects the risk-averse best response based on these beliefs (albeit possibly imperfectly). Two interrelated equations characterize game play under these assumptions. The first equation describes the expected utility from any given strategy, and the second describes how expected utility translates into behavior. When the row player in a matrix game has strategies  $A$  and  $B$  available and the column player decides between  $C$  and  $D$  (see Figure 1a), the quantal-response function asserts that the row player's probability of choosing  $A$  is an increasing function of the difference in expected utility (EU) between  $A$  and  $B$ , where the inverse of  $\eta_{\text{self}}$  governs the player's noisiness:

$$p(A) = \frac{1}{1 + e^{-\eta_{\text{self}}[EU(A) - EU(B)]}} \quad (1)$$

The expected utility of any given strategy, in turn, is given by computing the utility of the payoffs ( $x$ ), weighted by the row player's subjective belief,  $p^s$ , about whether the column player plays  $C$  or  $D$ :

$$EU(A) = p^s(C|k, \eta_{\text{other}}^s) U(x_{A,C}) + p^s(D|k, \eta_{\text{other}}^s) U(x_{A,D}), \quad (2)$$

where  $x_{i,j}$  represents the payoff for the row player when they choose row  $i$  and the column player chooses column  $j$ . In these equations, the key insights from the behavioral game theory literature appear in three different components. First, in assessing the expected utility from any given strategy, limited strategic sophistication (captured by  $k$ ) affects a player's expectation of their opponent's play and, hence, expected utility. Second, the expectation of the opponent's play also depends on subjective beliefs about the noisiness of the other player ( $\eta_{\text{other}}^s$ ). Following the literature, we initially assume that  $\eta_{\text{self}} = \eta_{\text{other}}^s \equiv \eta$ , meaning that each player believes that their opponent is as noisy as they are (McKelvey & Palfrey, 1995). Third, expected utility takes into account risk aversion, parameterized by constant absolute risk aversion,  $U(x) = (1 - e^{-\alpha x})/\alpha$ .

We quantify the performance of each model by benchmarking it against two extremes. First, to set a lower performance bound, we employed a random model, which predicts choice probabilities uniformly at random. Second, to set an upper performance bound, we employ a deep neural network (Multi-Layer Perceptron, MLP) that uses the game matrix as input and targets empirical choice frequencies as its output. The performance of all other models was assessed based on their *completeness*, which measures how well a model approximates the neural network upper bound from a starting point of random play (Fudenberg et al., 2022; Wright & Leyton-Brown, 2017). For example, a completeness of 50% indicates that a model's predictive accuracy is exactly in between those of the random model and the MLP. To determine overall model completeness, we averaged the completeness values calculated from two predictive accuracy metrics, MSE and  $R^2$ .

We found that allowing for limited strategic sophistication (level- $k$ ), quantal-response noise and risk aversion has large effects on model fit (see Figure 2e). The standard Nash equilibrium achieves only 24% completeness. A model that combines level-1 thinking, quantal-response noise and risk aversion achieves 82% completeness. While these

results illustrate that the success of behavioral game theory also translates into the much larger space of games that we analyzed, they also reveal substantial room for improvement of the structural decision-making model relative to the deep neural network. To close this gap, we considered what might be missing from existing theoretical accounts of strategic behavior.

A fundamental characteristic of existing models in the behavioral game theory literature is their context-invariance: they are uniformly applied and estimated, with identical parameters, irrespective of the characteristics of the game. However, evidence from related decision-making domains, such as choice between risky lotteries (Enke & Shubatt, 2023; Peterson et al., 2021), suggests that the parameters of behavioral models can be highly context-sensitive, for example because the complexity or cognitive difficulty of decisions varies across problems. The deep neural network model represents an extreme case of potential context-dependence because it is capable of forming its predictions based on the specific characteristics of each game. Yet while this model achieves high prediction accuracy, it is less useful for understanding human strategic decision-making because it is an uninterpretable “black box,” lacking the instructive format of structural decision-making models.

To build a bridge between these two approaches, we progressively introduced context-dependence into the structural decision-making models, by systematically substituting the structural parameters of behavioral models with a neural network that is responsive to game-specific features. We focus on three key behavioral primitives: (i) the level of strategic sophistication (i.e.,  $k$ ), (ii) the player’s level of noisiness (i.e.,  $\eta_{\text{self}}$ ; neural quantal response model), and (iii) the player’s beliefs regarding the noisiness of others (i.e.,  $\eta_{\text{other}}^s$ ; neural belief noise model). We allowed each of these structural parameters to be endogenously determined by the game matrices through a neural network, thus introducing context-dependence into the structural decision-making models in a disciplined and interpretable manner. To capture the context-dependence of  $k$  using neural networks, we let the neural network model the distribution of  $k$ :  $p(k)$ . As a result, the strategic choice predicted by this model can be interpreted as a mixture of players operating at different levels of  $k$  (See Supplementary Information for details).

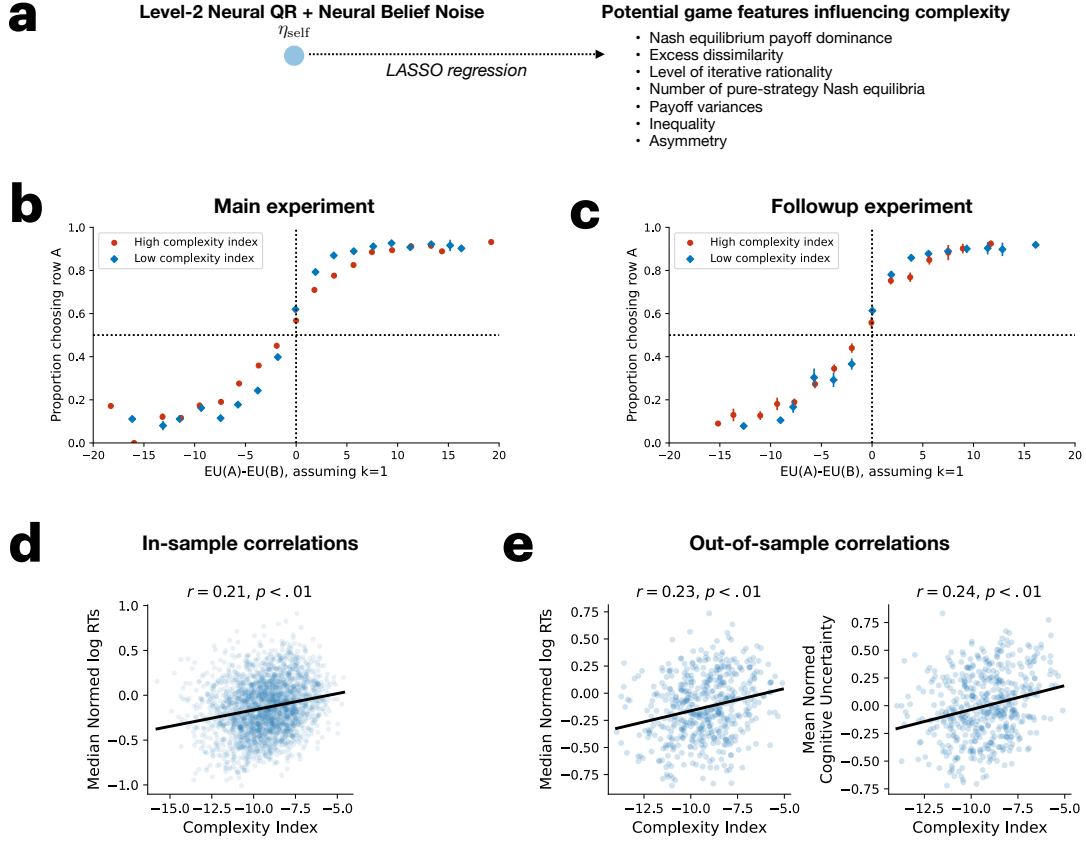
As an example, consider the level- $k$  neural quantal response model (Figure 2c), in which the player’s noisiness was predicted by a neural network. In this approach, an MLP was used to model the function  $\eta_{\text{self}} = f_{\text{MLP}}(\text{game matrix})$ , allowing the  $\eta_{\text{self}}$  parameter to vary across games. Once trained, the MLP can predict  $\eta_{\text{self}}$  for any given game matrix. By combining the strategic sophistication level  $k$  with the MLP-predicted  $\eta_{\text{self}}$ , the level- $k$  quantal-response function iteratively applies the quantal-response function  $k$  times to produce the probability of the row player choosing row  $A$  for a given game matrix. Similarly, for all other neurally-augmented behavioral models, the neural networks effectively learn a function that maps the game matrix to the parameters in the original behavioral model. Overall, the procedure of augmenting behavioral models with neural networks intuitively captures the possibility that players’ level of strategic sophistication or their noisiness is not fixed but, instead, depends on features of the game.

As illustrated in Figure 2e, integrating neural networks (i.e., MLPs) into the level- $k$  quantal-response model to account for context-dependence significantly enhances model completeness. When neural networks replaced all three structural parameters, model completeness reached 97%. The second-best model, achieving a completeness of 96%, maintained strategic sophistication at a fixed level of  $k = 2$ , yet allows for context-dependence in  $\eta_{\text{self}}$  and  $\eta_{\text{other}}^s$ . This context-dependence (i.e., game-specific noisiness) suggests an important role for decision-making difficulty – in some games, it is considerably easier to identify one’s best response than in others, even fixing a given level of strategic sophistication. Moreover, by comparing models that vary solely in  $\eta_{\text{self}}$  or  $\eta_{\text{other}}^s$  across different games (see Table S2), we found that variations or context dependence in  $\eta_{\text{self}}$  play a more significant role. This finding suggests that context dependence has a greater impact on individuals’ ability to optimize their own responses than on their ability to infer the likely actions of others.

To build intuition for why context-dependent noisiness has large effects on model completeness, Figure 3b shows the game-level link between the share of subjects choosing strategy  $A$  and the difference in expected utility between strategies  $A$  and  $B$ , as estimated from the level-2 neural quantal response model. We split the sample by the median estimated noisiness ( $\eta_{\text{self}}$ ). We see a strong attenuation pattern: for games with below-median noisiness, the link between strategy choice frequencies and expected utility estimates is considerably more compressed. Our level-2 neural quantal response model, a context-dependent model, achieves higher completeness both because it learns in which types of games players are more likely to best-respond, and because it learns in which types of games players are more likely to identify the best response of their opponent (given the presumed level of strategic sophistication,  $k$ ).

The high performance of our context-dependent models indicates that both the ability of players to optimally respond to their beliefs and their capacity to reason about others’ behaviors are contingent on the specific game being played. We attribute this sensitivity to *game complexity*, by which we mean the cognitive difficulty of (i) forming beliefs about other players’ strategies and (ii) optimally responding to these beliefs. To quantitatively define and validate a metric of game complexity, we first developed a composite index of game complexity, and then validated it by showing correlations with independent markers of complexity in a preregistered second experiment.

We leveraged our unusually large and diverse set of games to extract some of the specific game features that drive complexity, and aggregated them into a sparse and interpretable index of game complexity. To this effect, we analyzed the predicted  $\eta_{\text{self}}$  values from the context-dependent level-2 quantal response model (see Figure 3a: the Level-2 Neural QR + Neural Belief Noise model) and conducted a LASSO regression on a large set of structural game features (see Supplementary Information for detailed definitions). This analysis identified a concise set of influential game features. Some features are prominent in the literature, such as the number of steps of iterative reasoning required for equilibrium choices. Other features are more novel, including (i) a measure of the cognitive difficulty of navigating tradeoffs across different strategies, akin to work on dissimilarity and trade-

**Figure 3**

**Developing an interpretable complexity index for strategic games.** (a) To construct an index of game complexity, we use LASSO regressions to learn game features that correlate with the game-specific  $\eta_{self}$  parameter that is estimated by the MLP in the Level-2 Neural QR and Neural Belief Noise model. (b) The psychometric functions illustrate the relationship between the expected utility differences of two strategies and the proportion of choices for strategy A. Expected utility was calculated under the assumption of a level-1 player. Red (blue) dots represent high (low) complexity games, determined by a median split on the complexity index. Error bars represent  $\pm SE$ . (c) The same psychometric function and effect of complexity was found in the followup experiment. (d) The complexity index shows a statistically significant correlation with response times (RTs) in the games of the main experiment. (e) The complexity index generalized to the followup experiment and demonstrated statistically significant correlations with both RTs and cognitive uncertainty ratings.

off complexity in the literature on lottery choice (Enke & Shubatt, 2023; Shubatt & Yang, 2024); (ii) the variance and scale of payouts (Peterson et al., 2021); and (iii) the inequality and asymmetry in payouts between players.

We collapsed these interpretable game features into a composite index of complexity (see Supplementary Information). This index can be structurally interpreted as capturing the magnitude of the negated  $\eta_{\text{self}}$  predicted by game features. Because this index is defined based on objective game features, it can be readily computed by other researchers in any standard dataset.

A first piece of evidence that our complexity index indeed captures the difficulty of strategic decision-making is that, in our main experiment, the index has shown a positive correlation with response times (Pearson’s  $r = 0.21, p < .01$ , see Figure 3d), indicating that individuals tend to spend more time thinking in games that we classify as more complex.

To reinforce and broaden these within-sample correlations and our interpretation of the index, we conducted a preregistered follow-up experiment testing a new set of 500 games on a new cohort of participants. This experiment adhered closely to the procedure of the main experiment, with the sole modification being that, after each game, participants were required to report their cognitive uncertainty (in percentage terms) about whether the strategy they selected is actually their best decision (Enke & Graeber, 2023). The results confirmed that the index robustly predicts out-of-sample behavioral outcomes. As shown in Figure 3e, we replicated the positive correlation between RTs and the complexity index (Pearson’s  $r = 0.23, p < .01$ ). Additionally, we observed a positive correlation between cognitive uncertainty and the complexity index (Pearson’s  $r = 0.24, p < .01$ ), suggesting that participants tend to exhibit higher cognitive uncertainty in their strategic choices when faced with more complex games. Finally, in this follow-up experiment, we also studied the ability of our complexity index to predict behavioral attenuation in strategic choice. Figure 3c shows that the link between strategy choice frequencies and estimated expected utility differences is again considerably more compressed in the more complex problems. These findings confirm that our complexity index can generalize to out-of-sample strategic decisions.

Large-scale experiments and machine learning techniques have significantly aided our exploration of the vast space of cognitive mechanisms underlying strategic choices. Our findings reveal that a player’s own noisiness in responding to an opponent’s behavior, as well as their beliefs about the opponent’s noisiness, are critical factors in determining initial game play. Moreover, the degrees of noise vary across different games, displaying significant context dependence. To further understand the game features that contribute to this context dependence, we developed a complexity index that quantifies a player’s noisiness. This index is interpretable and readily generalizable to other matrix games, as it is based solely on features derived from the game matrix. The follow-up experiment confirmed that the complexity index effectively captures various aspects of human behavior in matrix games with differing levels of complexity. These results illustrate the promise of large-scale experiments and machine learning methods in furthering our understanding of strategic decision-making, in particular given the emerging body of theoretical work on complexity in behavioral game theory.

### Acknowledgments

This work and related results were made possible with the support of the NOMIS Foundation. We thank Nick Chater and Shengwu Li for helpful discussions.

### References

- Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1(4), 0064.
- Bell, R., & Cover, T. M. (1988). Game-theoretic optimal portfolios. *Management Science*, 34(6), 724–733.
- Bloom, P. (2012). Religion, morality, evolution. *Annual Review of Psychology*, 63, 179–199.
- Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., & Zhang, Q. (2018). *JAX: Composable transformations of Python+NumPy programs* (Version 0.3.13). <http://github.com/google/jax>
- Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton university press.
- Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3), 861–898.
- Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 69(5), 1193–1235.
- Crawford, V. P., Costa-Gomes, M. A., & Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51(1), 5–62.
- Devetag, G., & Warglien, M. (2008). Playing the wrong game: An experimental analysis of relational complexity and strategic misrepresentation. *Games and Economic Behavior*, 62(2), 364–382.
- Enke, B., & Graeber, T. (2023). Cognitive uncertainty. *The Quarterly Journal of Economics*, 138(4), 2021–2067.
- Enke, B., & Shubatt, C. (2023). *Quantifying lottery choice complexity* (tech. rep.). National Bureau of Economic Research.
- Eyster, E. (2019). Errors in strategic reasoning. *Handbook of Behavioral Economics: Applications and Foundations*, 2, 187–259.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868.
- Fudenberg, D., Kleinberg, J., Liang, A., & Mullainathan, S. (2022). Measuring the completeness of economic models. *Journal of Political Economy*, 130(4), 956–990.
- Fudenberg, D., & Liang, A. (2019). Predicting and understanding initial play. *American Economic Review*, 109(12), 4112–4141.
- Fudenberg, D., & Tirole, J. (1991). *Game theory*. MIT press.

- Gächter, S. (2004). Behavioral game theory. *Blackwell handbook of judgment and decision making*, 485–503.
- Golman, R., Bhatia, S., & Kane, P. B. (2020). The dual accumulator model of strategic deliberation and decision making. *Psychological Review*, 127(4), 477.
- Harsanyi, J. C. (1995). A new theory of equilibrium selection for games with complete information. *Games and Economic Behavior*, 8(1), 91–122.
- Hennigan, T., Cai, T., Norman, T., Martens, L., & Babuschkin, I. (2020). *Haiku: Sonnet for JAX* (Version 0.0.10). <http://github.com/deepmind/dm-haiku>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Mauersberger, F., & Nagel, R. (2018). Levels of reasoning in keynesian beauty contests: A generative framework. In *Handbook of computational economics* (pp. 541–634, Vol. 4). Elsevier.
- McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica: Journal of the Econometric Society*, 803–836.
- McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1), 6–38.
- Morrow, J. D. (1994). *Game theory for political scientists*. Princeton University Press.
- Murnighan, J. K., Roth, A. E., & Schoumaker, F. (1988). Risk aversion in bargaining: An experimental study. *Journal of Risk and Uncertainty*, 1(1), 101–124.
- Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5), 1313–1326.
- Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D., & Griffiths, T. L. (2021). Using large-scale experiments and machine learning to discover theories of human decision-making. *Science*, 372(6547), 1209–1214.
- Rawls, J. (1971). *A theory of justice*. Harvard University Press.
- Robinson, D., & Goforth, D. (2005). *The topology of the 2x2 games: A new periodic table* (Vol. 3). Psychology Press.
- Rosenthal, R. W. (1989). A bounded-rationality approach to the study of noncooperative games. *International Journal of Game Theory*, 18(3), 273–292.
- Schelling, T. C. (1980). *The strategy of conflict: With a new preface by the author*. Harvard University Press.
- Shubatt, C., & Yang, J. (2024). Similarity and comparison complexity. *arXiv preprint arXiv:2401.17578*.
- Stahl, D. O., & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization*, 25(3), 309–327.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11).
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... SciPy 1.0 Contributors. (2020). *SciPy 1.0: Fundamental Algorithms*

- for Scientific Computing in Python. *Nature Methods*, 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Von Neuman, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.
- Weizsäcker, G. (2003). Ignoring the rationality of others: Evidence from experimental normal-form games. *Games and Economic Behavior*, 44(1), 145–171.
- Wright, J. R., & Leyton-Brown, K. (2017). Predicting human behavior in unrepeated, simultaneous-move games. *Games and Economic Behavior*, 106, 16–37.



## Supplementary Information

### Materials and Methods

#### Main Experiment

##### *Game generation algorithm*

The games2p2k dataset comprises 2,416 instances of  $2 \times 2$  normal-form games and a total of 93,460 strategic choices made by 4,673 participants. Each game is uniquely identified by its payoff matrix, an 8-element vector encapsulating all payoff details for both row and column players. All payoffs are represented as integers and are restricted to a range of 1 to 50. Consequently, the set of all possible  $2 \times 2$  games encompasses  $50^8$  games before considering permutations of game matrices.

To generate games from this space, we employed a random generation process to create 8-item payoff matrices, where each item represents a random draw from a two-tiered uniform distribution  $U[1, u]$  where  $u \sim U[1, 50]$ . Next, we excluded games that lacked pure-strategy Nash equilibria and categorized each remaining game using Robinson and Goforth’s topology for  $2 \times 2$  games (Robinson & Goforth, 2005). This process results in a diverse set of 142 distinct game types. Given the prevalence of dominance games, which constituted 87.5% of the game types, we selectively reduced their representation in our dataset. Specifically, we generated 3 instances for each double-dominance game, 8 instances for each single-dominance game, and 22 instances for each non-dominance game, culminating in a comprehensive collection of 1,208 games. Given that each game is played by both the row and column players, our dataset comprises a total of 2,416 (i.e.,  $1,208 \times 2$ ) game instances.

##### *Participants*

We recruited a total of 4,900 participants via the Prolific Academic platform, out of which 4,673 individuals (1,942 males, 1,782 females, and 949 who opted not to disclose their gender) successfully completed the 10-minute experiment. These participants ranged in age from 18 to 86, with a median age of 37. Participants were required to be from the US and have completed at least 100 submissions with a 95% acceptance rate or higher on the Prolific Academic platform. Participants were guaranteed a base monetary compensation of \$2.00 (hourly rate of \$12.00), with the possibility of an additional bonus up to \$0.50 contingent on the result of a randomly selected game (1 point equals to \$0.01). The experimental sessions were carried out in Dec 2023. Informed consent was obtained from all participants (Princeton University IRB number 10859: “Computational Cognitive Science”, and Harvard University IRB number 16-1753).

##### *Procedure*

The experiment was programmed using Dallinger 9.11.0. Before the beginning of the main experiment, participants were provided with instructions and an example of

a normal-form game matrix. To ensure comprehensive understanding of the game representation, participants were required to correctly answer a pre-experiment multiple-choice quiz. Following this, participants were presented with a sequence of 20 one-shot games in normal form. All game presentations were adapted to a row player perspective, thereby making all participants choose between Row *A* and Row *B*, while the specifics of the game and the role they assumed were recorded in the background. Participants only played a game once. As the primary objective of our task was to observe choices in one-shot games, no practice or coaching period was provided.

To mitigate learning effects and preclude reputation building, each trial involved a different game and a new opponent for every participant. That is, participants were anonymously and randomly paired for each game. Feedback pertaining to a participant's performance in a specific game was withheld. The only occasion where feedback was provided occurred post-completion of the entire experimental session, indicating the bonus earned by the participant. The sequence in which the games were presented was randomly varied across participants. Moreover, for every game, the rows and the columns of the payoff matrix have equal chances being swapped, resulting in a total of 4 possible permutations for a game. The bonus was determined by randomly selecting one game played by the participant and their corresponding opponent in that game.

### **Follow-up Experiment**

This experiment was preregistered at <https://osf.io/xrvaw/>.

#### ***Games***

We employed the identical game generation algorithm used in the main experiment to create a set of 500 new normal-form games, maintaining similar proportions of dominance games as observed in the main experiment.

#### ***Participants***

We recruited another 1,013 participants from the Prolific Academic platform, of which 1,008 (346 males, 416 females, and 246 who opted not to disclose their gender) successfully completed the 10-minute experiment. Participant ages ranged from 18 to 88 years, with a median age of 35. The same filter used in the main experiment were applied here: participants were required to be from the U.S. and to have completed at least 100 submissions with a 95% acceptance rate or higher. As in the main experiment, participants received a base compensation of \$2.00 (hourly rate of \$12.00), with the possibility of earning a bonus of up to \$0.50 from a randomly selected game. The experimental sessions were conducted in April 2024. Informed consent was obtained from all participants (Princeton University IRB number 10859: "Computational Cognitive Science", and Harvard University IRB number 16-1753).

### ***Procedure***

The follow-up experiment replicated the exact procedure of the original study, with the sole modification that after participants made a strategic choice for a game, they were presented with a secondary question: “How certain are you that choosing [X] is actually your best decision?” where  $X$  denotes the selected option. Note that this query remained concealed until after a choice had been made. Participants were then required to express their confidence in their decision by adjusting a slider ranging from 0% (least confident) to 100% (most confident). The slider’s thumb was also hidden until the participant interacted with it by clicking on the slider bar.

### **Data Preprocessing**

We collected three types of behavioral measures from the two experiments (see Table S1 for an overview): strategic choices, response times (RTs) in making these choices, and subjective certainty ratings (confidence) associated with the choices. We aggregated data at the trial level to derive game-level behaviors. The empirical choice frequency, calculated across different participants for the same game, was used as the probability of choosing a particular strategy in that game. This measure served as the target variable for subsequent model fittings. RTs were initially log-transformed:  $\hat{RT} = \ln RT$ . Subsequently, we normalized these log-transformed RTs within participants by subtracting each participant’s mean  $\hat{RT}$  and dividing by the standard deviation of their  $\hat{RT}$ . The median of these normalized  $\hat{RT}$  values was then used to represent the game-level RTs. Confidence judgments were normalized within each participant, and the mean of these normalized confidence values was calculated to represent game-level confidence.

**Table S1**

*Overview of Behavioral Measures Collected in the Main and Follow-up Experiments.*

Behavioral measures	Statistical properties	Recording experiments
Choices	Binary outcomes	Both experiments
Response Times	Non-negative integer values	Both experiments
Confidences (Negated Cognitive Uncertainty)	Continuous range [0, 1]	Follow-up experiment

### **Model Details**

#### **Context-invariant Models**

The first class of models for human strategic choices comprises context-invariant models. These are parametric models with free parameters that remain constant across different games, hence the term “context invariance.” This consistency in parameters, however, does not imply that the models will predict identical behaviors for the same set of parameters across various games.

**Nash Equilibrium in Pure and Mixed Strategies.** In matrix games, a pure-strategy Nash equilibrium (PSNE) is an outcome where each player chooses a single action (or strategy) and no player can benefit by changing their action, given that the other players' actions remain the same. In other words, every player's choice is their best response to the choices of the other players. By contrast, a mixed-strategy Nash equilibrium (MSNE) occurs when players randomize over possible actions according to certain probabilities, and no player can improve their expected payoff by unilaterally changing their own probability distribution over actions. In MSNE, each player's mixed strategy is the best response to the mixed strategies of the other players.

**Level- $k$  Model.** The level- $k$  model in behavioral game theory explains strategic choices by assuming varying degrees of strategic sophistication among players (Costa-Gomes et al., 2001; Mauersberger & Nagel, 2018; Nagel, 1995). Level-0 players select their actions without strategic consideration, often randomly or according to a fixed rule, and they do not attempt to predict the actions of others. Level-1 players, assuming that all other players are at level 0, make decisions that are best responses to the expected actions of level-0 players. Each player at level  $k$  selects their action based on the belief that others are at level  $k - 1$ . Thus, the level- $k$  model captures a hierarchy of strategic reasoning, where each successive level represents a deeper anticipation of others' behavior based on the preceding level's actions.

**Quantal Response Equilibrium.** In a Nash equilibrium, each player's strategy is the best response to the strategies of the other players, and no player can improve their outcome by unilaterally changing their strategy. Conversely, the quantal response model offers a more realistic perspective by incorporating the idea that players may not always make perfectly rational decisions (McKelvey & Palfrey, 1995; Rosenthal, 1989). Instead of always choosing the best response, players select strategies with probabilities that increase with the expected payoff, introducing a noisy response. This means that players are more likely to choose better strategies, but they can still sometimes make suboptimal choices.

In the quantal response equilibrium (QRE), players are not assumed to be perfectly rational. Rather, they respond to expected payoffs in a probabilistic manner. The probability of selecting a particular strategy increases with the expected payoff of that strategy, yet suboptimal choices remain possible. Specifically, we used a logit quantal response function (McKelvey & Palfrey, 1995). Each player's strategy is influenced by the probability distribution of the other players' strategies, leading to a more flexible and realistic equilibrium concept.

**Level- $k$  Quantal Response Model.** While the level- $k$  model employs best response functions and the QRE model utilizes equilibrium solutions, the level- $k$  quantal response (QR) model integrates both approaches (Golman et al., 2020; Stahl & Wilson, 1994). In this model, level- $k$  players assume that all other players are at level  $k - 1$ , and they make decisions that are *quantal*, rather than best, responses to the expected actions of level- $k - 1$  players. Furthermore, it is important to note that level- $k$  players expect level- $k - 1$  players to apply the same quantal response logic to level- $k - 2$  players, continuing this pattern until reaching the random actions of level-0 players. In essence, level- $k$  players believe that

level- $k - 1$  players respond noisily to the actions of players one level below them, mirroring their own quantal response function.

**Level- $k$  Quantal Response and Belief Noise Model.** We also considered a simple extension of the Level- $k$  QR model by assuming that players might not believe their opponents exhibit the same level of noisiness as themselves, namely *Belief Noise*. There is empirical evidence supporting this assumption where, for example, people tend to underestimate the rationality of their opponents in strategic choices (Weizsäcker, 2003). Specifically, players maintain two sets of quantal response functions simultaneously: one for themselves when they need to respond quantally to their opponents’ choices, and another set reflecting what they believe their opponents use when responding to their own choices.

### Context-dependent Models

The other class of models exhibits context dependence, meaning that the parameter values of these models change in response to different game matrices. This context dependence is implemented using a Multilayer Perceptron (MLP) that processes game matrices as inputs and generates corresponding model parameters.

**Multilayer Perceptron.** The neural networks employed in this study use a consistent MLP architecture, featuring 3 hidden layers with 300 neurons each. Sigmoid functions are applied for all nonlinear activations, except when the network output is a probability, in which case a softmax function is used. The highest performance in modeling the games2p2k dataset was achieved by an MLP that directly predicts choice probabilities from the game matrix input:

$$p(A) = f_{MLP}(\text{game matrix}) \quad (\text{S1})$$

**Neural Network Augmented Behavioral Models.** We focus on adapting three key behavioral primitives using MLPs. First, we make the level of strategic sophistication (i.e.,  $k$ ) dependent on the game matrices. However, since levels are discrete and non-differentiable variables, we extend this concept by assuming a distribution  $p(k)$  of participants at each level  $k$ , where  $k = 0, 1, 2, 3$ . This approach effectively creates a mixture model that represents a range of Level- $k$  players. Next, the player’s own level of noisiness in response to opponents’ choices (i.e.,  $\eta_{\text{self}}$ ) and the player’s beliefs about the noisiness of their opponents (i.e.,  $\eta_{\text{other}}$ ) are modeled using two additional MLPs. The outputs of these MLPs are then input into a quantal response function to generate choice probabilities.

### Optimization and Cross-validation

We optimized all models by minimizing the mean squared errors (MSE) between the predicted outcomes and the actual proportions of  $A$  (or “up”) choices made by row players in each game. For models incorporating neural networks, they were implemented using the Python libraries `jax` (Bradbury et al., 2018) and `haiku` (Hennigan et al., 2020). The games2p2k dataset was partitioned randomly into training (80%), validation (10%), and testing (10%).

**Neural network training details.** Models involving neural networks were trained in batches of 64 games each. The dataset was also augmented through permutations. For each game, three additional duplicates were created by switching the rows, switching the columns, and switching both the rows and columns. The choice probabilities were then adjusted accordingly. Training was terminated when validation errors increased consecutively over two evaluation intervals of 100 epochs. The performance of the trained models was assessed on the testing set, with a particular emphasis on evaluating their generalization capabilities. The learning rate was fixed at  $10^{-3}$ , and the Adam optimizer was employed (Kingma & Ba, 2014).

**Context-invariant models training details.** Models without neural network components underwent optimization using the Nelder-Mead method via the `minimize` function of the Scipy library (Virtanen et al., 2020). Training these models involves randomly splitting the dataset into training (90%) and testing (10%) sets. The optimization phase concluded upon the convergence of the Nelder-Mead algorithm on the training set. Model performance was also evaluated on the testing set.

The cross-validation for both context-invariant and context-dependent models was repeated a total of 10 rounds for all models. Different random partition over games was performed in each round. Both the mean and standard error for the MSE and the  $R^2$  were recorded for the testing set (see Table S2).

### Developing a Game Complexity Index

In our computational modeling, we identified the player’s level of noisiness (i.e., the  $\eta_{\text{self}}$  parameter from context-dependent models) as an effective target for deriving an interpretable index that reflects perceptions of game complexity. This choice is based on the premise that game complexity is likely linked to deviations from rational behavior, which are captured by the  $\eta_{\text{self}}$  parameter. To develop this index, we employed a two-step procedure: First, we reviewed existing game features known in the literature and introduced a series of novel features. This process created a comprehensive set of candidate game features that potentially elucidate aspects of game complexity. Subsequently, we refined this set to identify the most informative features by employing LASSO regression, using the  $\eta_{\text{self}}$  parameter as the dependent variable and the full suite of game features as independent variables.

### Game Topology

Using the illustrative game depicted in Figure S1, we summarize Robinson and Goforth’s topology for  $2 \times 2$  matrix games (Robinson & Goforth, 2005). Since both row and column players have 12 unique order graphs in the payoff space of their respective payoff matrices, this topology is constructed on a set of  $12 \times 12 = 144$  ordinal order graphs. For instance, from the perspective of a row player, the 12 order graphs including Chicken (or Hawk-Dove, Snowdrift,  $c > a > b > d$ ), Leader (or Battle of the sexes,  $c > b > a > d$ ), Hero ( $c > b > d > a$ ), Compromise ( $c > d > b > a$ ), Deadlock (or Altruist’s Dilemma,

		Col player (Other)	
		C	D
Row player (Self)	A	<b>a</b>   <b>x</b>	<b>b</b>   <b>y</b>
	B	<b>c</b>   <b>z</b>	<b>d</b>   <b>w</b>

**Figure S1**

*Example of a game matrix: The payoff matrix for the row player ('self') is in blue, while the payoff matrix for the column player ('other') is in red.*

$c > d > a > b$ ), Prisoner's Dilemma ( $c > a > d > b$ ), Stag Hunt (or Trust,  $a > c > d > b$ ), Assurance ( $a > d > c > b$ ), Safe Coordination ( $a > d > b > c$ ), Peace (or Club,  $a > b > d > c$ ), Harmony ( $a > b > c > d$ ), Concord ( $a > c > b > d$ ). The same topology was applied to the column player's payoff matrix.

### Game Features

In this section, we provide a detailed description, along with examples, of the interpretable features of a game matrix. We also discuss their expected correlation with game complexity.

**Dominant Solvability.** Dominant solvable games, which means that a game is solvable by iterated deletion of dominated action, are generally considered easier to play because the payouts from one strategy consistently outperform all others. Players' abilities to identify the optimal strategy only require the capacity to best respond. This feature can be independently defined for both the player (row player) and the opponent (column player). For instance, considering the illustrative game depicted in Figure S1, we define the dominant solvability for both players (equality can occur in only one of these inequalities):

$$\text{DominantSolvable}_{\text{self}} = \begin{cases} 1 & \text{if } a \geq c \text{ and } b \geq d \text{ and not } (a = c \text{ and } b = d) \\ 1 & \text{if } a \leq c \text{ and } b \leq d \text{ and not } (a = c \text{ and } b = d) \\ 0 & \text{otherwise} \end{cases} \quad (\text{S2})$$

$$\text{DominantSolvable}_{\text{other}} = \begin{cases} 1 & \text{if } x \geq y \text{ and } z \geq w \text{ and not } (x = y \text{ and } z = w) \\ 1 & \text{if } x \leq y \text{ and } z \leq w \text{ and not } (x = y \text{ and } z = w) \\ 0 & \text{otherwise} \end{cases} \quad (\text{S3})$$

**Excess Dissimilarity.** Excess dissimilarity, initially introduced in (Enke & Shubbatt, 2023) to understand complexity in risky choices, quantifies the difficulty of aggregating values across states. In risky choices, the measure calculates payout differences between states for two lotteries, then deducts the difference in expected values. We expand this measure to two-player matrix games, further assuming that the opponent's strategies are random. Using the game illustrated in Figure S1, we compute the excess dissimilarity indices for both players as follows:

$$\text{Dissimilarity}_{\text{self}} = \frac{|a-c|}{2} + \frac{|b-d|}{2} - |\mu_{\text{up}} - \mu_{\text{down}}| \quad (\text{S4})$$

$$\text{Dissimilarity}_{\text{other}} = \frac{|x-y|}{2} + \frac{|z-w|}{2} - |\mu_{\text{left}} - \mu_{\text{right}}| \quad (\text{S5})$$

where  $\mu_{\text{up}} = \frac{a+b}{2}$ ,  $\mu_{\text{down}} = \frac{c+d}{2}$ ,  $\mu_{\text{left}} = \frac{x+z}{2}$ ,  $\mu_{\text{right}} = \frac{y+w}{2}$ .

**Levels of Iterative Rationality.** We define the levels of iterative rationality as the minimum level  $k$  at which the best-response dynamic converges to a fixed strategy. For  $2 \times 2$  games, we set the upper limit of  $k$  to 3 due to the possibility of infinite, cyclical best-response dynamics.

**Number of Nash Equilibria.** We analyze pure-strategy and mixed-strategy Nash equilibria separately (PSNE and MSNE) and determine the number of equilibrium strategies for a game matrix.

**Nash Equilibrium Payoff Dominance.** We compute whether each pure-strategy Nash equilibrium (PSNE) in a game yields the maximum possible payouts for both players. If at least one such equilibrium exists, we classify the game as exhibiting Nash equilibrium payoff dominance; otherwise, this feature is assigned a value of zero. Moreover, we compute Non-Nash Equilibrium Payoff Dominance, which is defined as the presence of an outcome that payoff-dominates all other outcomes but is not a PSNE.

**Nash Equilibrium Pareto Dominance.** Following (Fudenberg & Liang, 2019), we also assess whether a PSNE offers superior payouts compared to all other PSNEs. If such a PSNE exists, we classify the game as demonstrating Nash equilibrium Pareto dominance. If the game possesses only a single PSNE, it is automatically categorized as Pareto dominant. All other cases are classified as false.

**Pure Motives.** The notion of pure motives was examined in (Devetag & Warglien, 2008). In games governed by pure motives, players' choice preferences are expected to exhibit perfect rank correlation. This includes scenarios with positive correlations, such as coordination games, as well as negative correlations, like zero-sum games. Games featuring both antagonistic and coordination motives, such as the chicken game or the prisoner's dilemma, are said to be mixed motives games.

**Max Payouts.** For a given game matrix, we calculate the maximum potential payouts for each player. Referring to the example game in Figure S1, we define the maximum payouts as follows:

$$\text{Max}_{\text{self}} = \max\{a, b, c, d\} \quad (\text{S6})$$

$$\text{Max}_{\text{other}} = \max\{x, y, z, w\} \quad (\text{S7})$$



**Payoff Variances.** Variations in payouts may influence strategy selection. Here, we employ the canonical variance measure. To illustrate, let's examine the game matrix in Figure S1:

$$\text{PayoffVar}_{\text{self}} = \frac{1}{4} \left[ (a - \mu_{\text{up}})^2 + (b - \mu_{\text{up}})^2 + (c - \mu_{\text{down}})^2 + (d - \mu_{\text{down}})^2 \right] \quad (\text{S8})$$

$$\text{PayoffVar}_{\text{other}} = \frac{1}{4} \left[ (x - \mu_{\text{left}})^2 + (z - \mu_{\text{left}})^2 + (y - \mu_{\text{right}})^2 + (w - \mu_{\text{right}})^2 \right] \quad (\text{S9})$$

**Deviations from Zero-Sum Games.** Zero-sum games exhibit a distinct payout structure where gains for one player directly correspond to losses for another. To gauge deviations from a standard zero-sum game, we introduce a measure. First, we compute the absolute sum of payout differences for each player when jointly deviating from an outcome. Subsequently, we aggregate these absolute sums across all four possible deviations within a  $2 \times 2$  game. Illustrated through the example game in Figure S1, we calculate this metric as outlined below:

$$\text{NonZeroSum} = |a - c + x - z| + |a - b + x - y| + |c - d + z - w| + |b - d + y - w| \quad (\text{S10})$$

**Inequality in Payouts.** Social preferences may also play a role in shaping strategic decisions (Fehr & Schmidt, 1999). To capture specific aspects of these preferences, the following features have been devised. Inequality in payouts assesses the disparity in the maximum potential payouts players could attain from the game. In the context of the game depicted in Figure S1, this can be computed as follows:

$$\text{Inequality} = \max\{a, b, c, d\} - \max\{x, y, z, w\} \quad (\text{S11})$$

**Asymmetry in Payouts.** Another source of payout inequality arises from variations in payouts when both players employing a similar strategy. This can be quantified as the extent of deviations from a symmetric game, where payouts would be identical if both players adopted the same strategy. To compute the asymmetry in payouts for the game depicted in Figure S1, we proceed as follows:

$$\text{Asymmetry} = \frac{1}{4} \left[ |a - x| + |b - z| + |c - y| + |d - w| \right] \quad (\text{S12})$$

### Expected Relationships Between Game Features and Complexity

We have now developed a set of game features that characterize a game matrix and are relevant to game complexity. Here, we present their predicted correlations with complexity and provide justifications for these expectations (see Table S3). Broadly, these justifications can be summarized into the following categories.

First, games with strategies that players will never choose, as long as they are best responding to *any* belief, are considered less complex. Conversely, other types of games require rational choices that ensure both consistency in beliefs about opponents' strategies

and the ability to best respond to those beliefs. Dominant Solvability is the feature with this property.

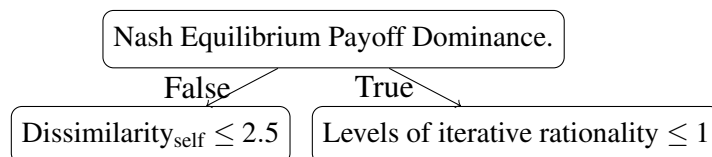
Second, game features that facilitate coordination are expected to reduce game complexity. Efficiency, defined by payoff dominance or Pareto dominance (Harsanyi, 1995; Schelling, 1980), is considered separately from rational equilibrium solutions. When equilibrium solutions coincide with efficient solutions, players can coordinate on the overlapping outcome, making the game easier to solve. Conversely, when different solution concepts do not align, the game becomes more challenging. Game features with this property include Nash Equilibrium Payoff and Pareto Dominance, Maximum Payouts, and Non-Nash Equilibrium Payoff Dominance.

Third, a game is generally expected to be less complex if players can easily adopt their opponent's perspective (Schelling, 1980). This occurs when players' preferences are perfectly rank-correlated, either positively (e.g., coordination games) or negatively (e.g., zero-sum games). The following game features capture this property: Pure Motives, Deviation from Zero-Sum Games, Inequality in Payouts, and Asymmetry in Payouts.

Fourth, games that require greater cognitive resources to process are intuitively more complex. For instance, choosing among multiple equilibria increases complexity, as reflected by the number of Pure Strategy Nash Equilibria (PSNE) and Mixed Strategy Nash Equilibria (MSNE). Moreover, the computation of expected values can vary in complexity, leading to the inclusion of features such as Excess Dissimilarity and Payoff Variance. Finally, iterative thinking steps needed to form consistent beliefs and choose an equilibrium strategy should also matter for players with bounded cognitive resources. This concept is captured by the Levels of Iterative Rationality.

### Selection of Game Features

**Decision tree regressions.** To identify important features, we applied decision tree regression using interpretable game features directly on the choice probabilities in games2p2k. By fixing the maximal tree depth to 3, we found that the most important feature was whether the game had a Nash Equilibrium that payoff-dominates all other outcomes. This was followed by excess dissimilarity of the player's own payoff matrix and levels of iterative rationality, which appeared in the second layer of the decision tree. The third layer was omitted as it contained repetitions of the features in the second layer, suggesting that these three features are critical in deciding strategic choices.



**LASSO regressions.** Considering the array of game features that may impact strategic decisions, we develop a game complexity index to explain the noisiness in players' strategic choices. This index is designed to encapsulate deviations from rationality

observed within our games. First, we normalized the game features across different games to ensure comparability among game features. Subsequently, we implemented a LASSO regression by setting the multiplier of the  $L_1$  term at 0.2, targeting the negated  $\eta_{\text{self}}$  parameter from our best-fitting level-2 quantal response model, with the normalized game features serving as independent variables. Results were summarized in Table S4.

### Robustness Check

To further validate our modeling results, we trained both context-invariant and context-dependent models using data from the main experiment and tested their performance using data from the follow-up experiment. For the context-invariant models, training was conducted on the entire main experiment dataset. For the context-dependent models, we applied an optional stopping rule to terminate neural network training by partitioning the main experiment data into training and validation sets with a 90/10 split. The remaining training details were consistent with those previously described. As shown in Table S5, our main results were successfully replicated.

### Experimental Instructions and Comprehension Checks

**Fig. S1.** Screenshot of experimental instructions presented to participants in the main experiment.

**Fig. S2.** Screenshot of the comprehensive check in the main experiment.

**Fig. S3.** Screenshot of a trial showing the game matrix to a participant in the main experiment.

**Fig. S4.** Screenshot of experimental instructions presented to participants in the follow-up experiment.

**Fig. S5.** Screenshot of a trial showing the game matrix and confidence rating to a participant in the follow-up experiment.

## Instructions

In this experiment, you will play a series of simple games in which you and another person each make a choice and the payoff each of you receive depends on both choices

In total, you will participate in **20** games. At the beginning of each game, you will be randomly matched with a different person who is also participating in this experiment on Prolific. You and the other person will then see a game displayed in a table such as the one below.

	Other's choice C	Other's choice D
Your choice A	1   5	2   3
Your choice B	3   4	4   2

The rows of the table show **your** options, and the columns the **other** player's options.

The payoffs you receive are shown in **blue** on the left hand side of the line in each cell of the table, whereas the other player's choice options and payoffs are shown in **red** on the right hand side of the line. The table is read as follows:

- If you choose A and the other player chooses C, you receive **1** point
- If you choose A and the other player chooses D, you receive **2** points
- If you choose B and the other player chooses C, you receive **3** points
- If you choose B and the other player chooses D, you receive **4** points

Note that the other player is in a similar situation as you are. This means that:

- If the other player chooses C and you choose A, the other player receives **5** points
- If the other player chooses C and you choose B, the other player receives **4** points
- If the other player chooses D and you choose A, the other player receives **3** points
- If the other player chooses D and you choose B, the other player receives **2** points

You will not be informed of the choice or the identity of the other player. All your decisions will be treated confidentially.

1 point = \$0.01

Keep in mind that you will be randomly rematched with a different player for every new game.

Your earnings in the experiment will be the payoffs obtained in a randomly selected game plus the \$2 participation fee

[Start pre-experiment quiz](#)

**Table S2**

*Comparisons Between Context-invariant and Neurally-Augmented Context-dependent Models*

Model	MSE (SE) ↓	$R^2$ (SE) ↑	Completeness (%) ↑
Random	.0875	.0003	0 (lower bound)
Nash	.1625	.2234	24
L1+QR	.0255 (.0006)	.7134 (.009)	77.5
L1+QR+Risk	.0218 (.0005)	.7509 (.0092)	82
L2+QR	.026 (.0006)	.6968 (.0077)	76.5
L2+QR+Risk	.0244 (.0005)	.7306 (.0056)	79
L2+QR+Belief+Risk	.0182 (.0005)	.7949 (.0062)	86
L3+QR	.0275 (.0006)	.6868 (.0072)	75
L3+QR+Risk	.0238 (.0005)	.7303 (.006)	79
L3+QR+Belief+Risk	.0181 (.0006)	.7906 (.0063)	86.5
QRE	.0269 (.0007)	.7112 (.0099)	76.5
QRE+Risk	.0251 (.0005)	.7213 (.0061)	78
QRE+Belief+Risk	.0183 (.0005)	.7888 (.006)	86
L1+QR	.0229 (.0011)	.773 (.0091)	82.5
L1+QR+Risk	.0145 (.0006)	.8361 (.0077)	91
L2+QR	.0206 (.0012)	.78 (.0095)	84
L2+QR+Risk	.0178 (.0006)	.7968 (.0065)	87
L2+QR+Belief+Risk	.0125 (.0003)	.862 (.0045)	94
L2+QR+Belief+Risk	.0291 (.0006)	.6895 (.0073)	74
L2+QR+Belief+Risk	.0109 (.0003)	.8784 (.0045)	96
L3+QR	.0389 (.0089)	.6944 (.0259)	68.5
L3+QR+Risk	.0187 (.0003)	.7856 (.0036)	85.5
L3+QR+Belief+Risk	.0124 (.0002)	.8615 (.0024)	94
L3+QR+Belief+Risk	.032 (.001)	.6776 (.0091)	71.5
L3+QR+Belief+Risk	.0119 (.0003)	.8674 (.0045)	94
QRE	.0245 (.0011)	.7492 (.0096)	80
QRE+Risk	.0188 (.0005)	.7904 (.0067)	86
QRE+Belief+Risk	.0125 (.0005)	.8604 (.0058)	94
QRE+Belief+Risk	.0282 (.0009)	.7231 (.0081)	76.5
QRE+Belief+Risk	.0116 (.0004)	.8732 (.0036)	95
L+QR+Risk	.0194 (.0003)	.7834 (.0052)	85
L+QR+Belief+Risk	.0096 (.0003)	.8931 (.0039)	97
MLP	.0073 (.0001)	.9194 (.0022)	100 (upper bound)

*Note.* The model completeness was determined by averaging the completeness values calculated from Mean Squared Error (MSE) and  $R^2$ , except for the Nash equilibrium model, where completeness was solely based on  $R^2$ . Colorized texts denote components of the model that are implemented as neural networks. Specifically, **QR** indicates that  $\eta_{\text{self}}$  is a neural network, **Belief** signifies that  $\eta_{\text{other}}^s$  is a neural network, and **L** means that a neural network predicts a mixture of level- $k$  players. The numbers in parentheses represent the standard errors from the 10-fold cross-validation.

**Table S3**

*Developed game features, both existing and novel, and their expected correlation with game complexity.*

Game feature	Expected correlation with game complexity
DominantSolvable <sub>self</sub>	-
DominantSolvable <sub>other</sub>	-
Nash Equilibrium Payoff Dominance	-
Nash Equilibrium Pareto Dominance	-
Pure Motives	-
Max <sub>self</sub>	-
Max <sub>other</sub>	-
Deviation from Zero-Sum Games	-
Levels of Iterative Rationality	+
Dissimilarity <sub>self</sub>	+
Dissimilarity <sub>other</sub>	+
Number of PSNE	+
Number of MSNE	+
Non-Nash Equilibrium Payoff Dominance	+
PayoffVar <sub>self</sub>	+
PayoffVar <sub>other</sub>	+
Inequality in Payouts	+
Asymmetry in Payouts	+

*Note.* Negative correlations suggest that the presence or higher values of certain game features correspond to a decrease in game complexity, while positive correlations indicate the opposite.

**Table S4***The results of LASSO regressions.*

Game feature	$-\eta_{\text{self}}$	$-\eta_{\text{other}}^s$
DominantSolvable <sub>self</sub>	.	.
DominantSolvable <sub>other</sub>	.	.
Dissimilarity <sub>self</sub>	0.28	.
Dissimilarity <sub>other</sub>	.	0.08
LevelIterRational	0.38	.
NumPSNE	.	.
NumMSNE	.	.
PayoffDomEquilibrium	-0.80	.
PayoffDomNonEquilibrium	.	.
ParetoDomEquilibrium	.	.
PureMotives	.	.
Max <sub>self</sub>	0.30	.
Max <sub>other</sub>	.	0.01
PayoffVar <sub>self</sub>	0.40	.
PayoffVar <sub>other</sub>	.	.
NonZeroSum	.	.
Inequality	0.85	.
Asymmetry	-0.09	.
Intercept	-9.28	-1.09
$R^2$	0.33	0.02

**Table S5**

*Comparisons of models trained on the main experiment and tested on the follow-up experiment.*

Model	MSE (SE) ↓	$R^2$ (SE) ↑
Nash	.1680	.2561
L1+QR+Risk	.0203	.758
L2+QR+Belief+Risk	.016	.8113
L3+QR+Belief+Risk	.0167	.8026
QRE+Belief+Risk	.0166	.8039
L1+QR+Risk	.0128 (.0001)	.8547 (.0032)
L2+QR+Belief+Risk	.0109 (.0002)	.875 (.0027)
L2+QR+Belief+Risk	.0102 (.0001)	.8868 (.0018)
L3+QR+Belief+Risk	.0108 (.0002)	.8798 (.0022)
L3+QR+Belief+Risk	.0104 (.0001)	.8855 (.0018)
QRE+Belief+Risk	.0111 (.0002)	.877 (.0029)
QRE+Belief+Risk	.0105 (.0001)	.8835 (.0018)
L+QR+Belief+Risk	.0081 (.0002)	.9065 (.0023)
MLP	.0073 (.0001)	.9194 (.0022)

*Note.* Colorized texts denote components of the model that are implemented as neural networks. Specifically, **QR** indicates that  $\eta_{\text{self}}$  is a neural network, **Belief** signifies that  $\eta_{\text{other}}^s$  is a neural network, and **L** means that a neural network predicts a mixture of level- $k$  players. The numbers in parentheses represent the standard errors from the 10-fold cross-validation.



## Pre-experiment quiz

In the main experiment, you will play a total of **20** games, each with a different person.

Now suppose that you are presented with the following game table:

	Other's choice C	Other's choice D
Your choice A	1   5	2   3
Your choice B	3   4	4   2

What determines your payoff?

My choice and the other player's choice ▼

In which color is your payoff shown in the table?

Blue ▼

If you choose A and the other player chooses C, what is your payoff?

1 ▼

If you choose B and the other player chooses D, what is the other player's payoff?

2 ▼

←Review instructions

Begin experiment→

	Other's choice C	Other's choice D
Your choice A	17   20	30   7
Your choice B	34   2	15   2

A new game has been created, and you are paired with a different player.

Please choose a row:

Choice A

Choice B

## Instructions

In this experiment, you will play a series of simple games in which you and another person each make a choice and the payoff each of you receive depends on both choices

In total, you will participate in **20** games. At the beginning of each game, you will be randomly matched with a different person who is also participating in this experiment on Prolific. You and the other person will then see a game displayed in a table such as the one below.

	Other's choice C	Other's choice D
Your choice A	1   5	2   3
Your choice B	3   4	4   2

The rows of the table show **your** options, and the columns the **other** player's options.

The payoffs you receive are shown in **blue** on the left hand side of the line in each cell of the table, whereas the other player's choice options and payoffs are shown in **red** on the right hand side of the line. The table is read as follows:

- If you choose A and the other player chooses C, you receive **1** point
- If you choose A and the other player chooses D, you receive **2** points
- If you choose B and the other player chooses C, you receive **3** points
- If you choose B and the other player chooses D, you receive **4** points

Note that the other player is in a similar situation as you are. This means that:

- If the other player chooses C and you choose A, the other player receives **5** points
- If the other player chooses C and you choose B, the other player receives **4** points
- If the other player chooses D and you choose A, the other player receives **3** points
- If the other player chooses D and you choose B, the other player receives **2** points

For each game, you will make two decisions:

- You will decide whether to play row A or row B
- We will ask you how certain you are about your decision. Specifically, we are interested in how likely you think it is (in percent) that the decision you took is actually your best decision, given your preferences and the available information

You will not be informed of the choice or the identity of the other player. All your decisions will be treated confidentially.

1 point = \$0.01

Keep in mind that you will be randomly rematched with a different player for every new game.

Your earnings in the experiment will be the payoffs obtained in a randomly selected game plus the \$2 participation fee

[Start pre-experiment quiz](#)

	Other's choice C	Other's choice D
Your choice A	5   14	6   19
Your choice B	17   13	6   9

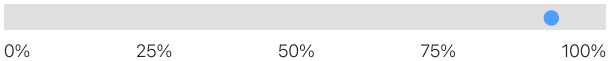
A new game has been created, and you are paired with a different player.

Please choose a row:

Choice A

Choice B

How certain are you that choosing **Row B** is actually your best decision?



Submit