nature human behaviour

Article

https://doi.org/10.1038/s41562-025-02230-5

Capturing the complexity of human strategic decision-making with machine learning

Received: 24 August 2024

Accepted: 28 April 2025

Published online: 25 June 2025

Check for updates

Jian-Qiao Zhu[1] ✉, Joshua C. Peterson[2], Benjamin Enke[3,4] & Thomas L. Griffiths[1,5]

Strategic decision-making is a crucial component of human interaction. Here we conduct a large-scale study of strategic decision-making in the context of initial play in two-player matrix games, analysing 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 this dataset predicts human choices with greater accuracy than leading theories of strategic behaviour, revealing systematic variation unexplained by existing models. By modifying this network, we develop an interpretable behavioural model that uncovers key insights: individuals' abilities to respond optimally and reason about others' actions are highly context dependent, influenced by the complexity of the game matrices. Our findings illustrate the potential of machine learning as a tool for generating new theoretical insights into complex human behaviours.

The classical 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 analysing matrix games, research has shown that human players often violate both of these assumptions[1–4]. Consequently, the effectiveness of these equilibria in explaining people's strategic behaviours is limited[1,5]. This has prompted the development of behavioural game theory, which has identified various extensions and refinements that produce a closer match to human decisions[6–9].

Despite a proliferation of behavioural 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[8,10]. 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 behaviour is limited. To explore these questions, we conducted a large-scale study of strategic behaviour by densely sampling the enormous space of 2 × 2 game structures. We use the resulting dataset to assess the explanatory power of leading behavioural 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 behaviour almost as well as the machine learning algorithm.

We procedurally generated a dataset of 2,416 matrix games involving monetary gains (Fig. 1b), substantially 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[10]). To systematically sample game matrices, our game generation algorithm is based on the Robinson and Goforth topology for 2 × 2 games, which is constructed using ordinal order graphs of payoffs[11]. 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 × 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 the Methods for details). To investigate human behaviour in these settings, we recruited 4,900 participants via Prolific, each of whom was instructed to participate in 20 distinct games, sampled randomly without replacement from the pool of procedurally generated games. In total, 93,460 strategic decisions were recorded. No feedback was provided to participants

[1]Department of Computer Science, Princeton University, Princeton, NJ, USA. [2]Faculty of Computing and Data Science, Boston University, Boston, MA, USA. [3]Department of Economics, Harvard University, Cambridge, MA, USA. [4]National Bureau of Economic Research, Cambridge, MA, USA. [5]Department of Psychology, Princeton University, Princeton, NJ, USA. ✉e-mail: jz5204@princeton.edu

 https://doi.org/10.1038/s41562-025-02230-5

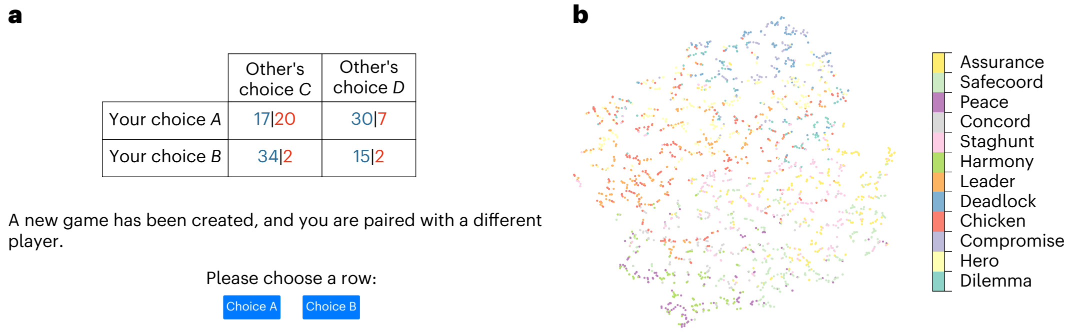

Fig. 1 | Matrix games. a, An example game interface presented to participants, who acted as the row player in each 2 × 2 game. The blue numbers represent the payoffs for the row player, who chooses between strategies *A* and *B*, while the red numbers represent the payoffs for the column player, who chooses between strategies *C* and *D*. **b**, Visualization of game space. Each game is uniquely represented by an eight-integer vector, corresponding to the payoffs to the two players under different configurations of choices. We used *t*-distributed stochastic neighbour embeddings[26] 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 colours represent the game topology specific to the row player following ref. 11.

between games, and the players were randomly rematched after each game. Thus, the observed behaviours 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 between model predictions and empirical choice frequencies at the game level on the training set, and then assess their performance in a validation set (Supplementary Information section B).

The models are based on three key insights from the behavioural 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[12,13]. The best-response function defines a player's optimal strategy based on their beliefs about their opponents' strategies in the game. Second, a player may exhibit noise in their decision-making and also take into account the noisiness of others' behaviour, as in quantal-response (QR) equilibrium models[1]. Third, players may be risk averse, meaning they prefer less uncertainty[8,14].

All models that we estimate are variants of our baseline behavioural model, which is a risk-averse level-*k* QR model (Fig. 2a)[10]. According to this model, a player forms beliefs about their opponent's behaviour (albeit possibly imperfectly) and selects the risk-averse best response based on these beliefs (albeit possibly imperfectly). The QR equilibrium model is not nested within the quantal level-*k* framework, as taking the limit as $k \to \infty$ does not necessarily lead to a QR equilibrium. Two interrelated equations characterize game play under these assumptions. The first equation describes how expected utility translates into behaviour, and the second describes the expected utility from any given strategy. When the row player in a matrix game has strategies *A* and *B* available and the column player decides between *C* and *D* (Fig. 1a), the QR 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 η_{self} governs the player's noisiness:

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

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*:

$$\mathrm{EU}(A) = p^{\mathrm{s}}\left(C|k, \eta^{\mathrm{s}}_{\text{other}}\right) U(x_{A,C}) + p^{\mathrm{s}}\left(D|k, \eta^{\mathrm{s}}_{\text{other}}\right) U(x_{A,D}), \qquad (2)$$

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 behavioural 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 (η^{s}_{other}). Following the literature, we initially assume that $\eta_{\text{self}} = \eta^{s}_{\text{other}} \equiv \eta$, meaning that each player believes that their opponent is as noisy as they are[1]. Third, expected utility takes into account risk aversion, parameterized by constant absolute risk aversion, $U(x) = (1 - \mathrm{e}^{-\alpha x})/\alpha$.

Results

We quantify the performance of each model by benchmarking it against two extremes. First, to set a lower performance bound, we used a random model, which predicts choice probabilities uniformly at random. Second, to set an upper performance bound, we use a deep neural network (multilayer 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[10,15]. Specifically, the completeness of a model compares the performance improvement of that model over the random model (measured in both mean squared error and R^2) with the corresponding improvement of the MLP model over the same baseline (see equation (B4) in the Supplementary Information for the detailed formulation). For example, a completeness score of 50% indicates that a model has achieved half of the improvement in predictive accuracy that the MLP model demonstrates over the random model.

We found that allowing for limited strategic sophistication (level-*k*), QR noise and risk aversion has large effects on model fit (Fig. 2e). The standard Nash equilibrium achieves only 22% completeness. A model that combines level-1 thinking, QR noise and risk aversion achieves 82% completeness. While these results illustrate that the success of behavioural game theory also translates into the much larger space of games that we analysed, 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 behaviour.

A fundamental characteristic of existing models in the behavioural 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

 https://doi.org/10.1038/s41562-025-02230-5

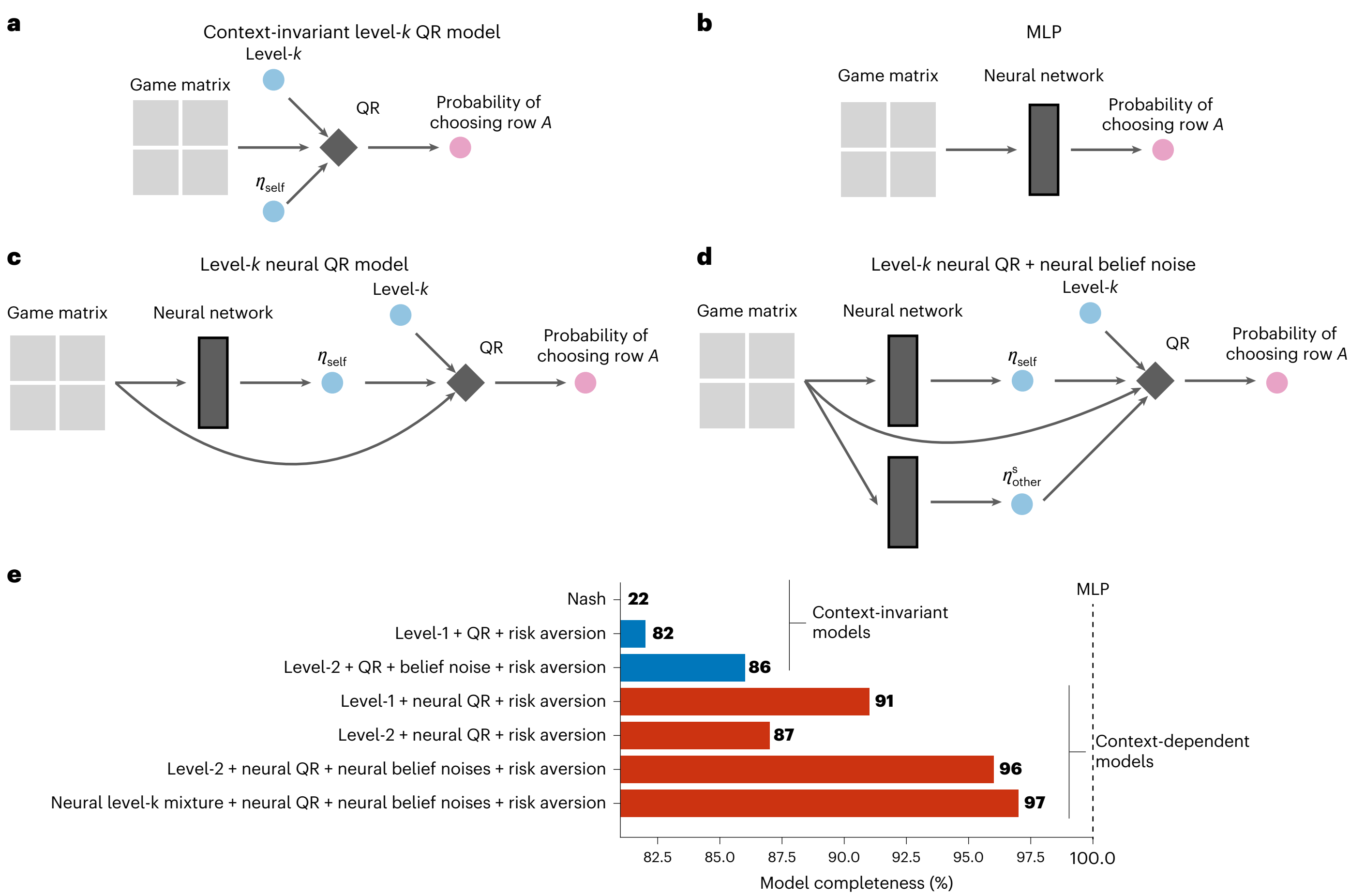

Fig. 2 | Model comparisons. a, The context-invariant level-*k* QR model involves three parameters that do not vary across games: strategic sophistication (that is, *k*), the players' noisiness (that is, η_{self}) and risk aversion. **b**, A MLP model directly uses the game matrix as input to estimate choice probabilities, without imposing any specific game-theoretic decision-making structure. **c**, The level-*k* neural QR model is a context-dependent model allowing the η_{self} parameter to vary across games. It uses an MLP model, which uses the game matrix as input to estimate game-specific η_{self}. **d**, The level-*k* neural QR and neural belief noise model extends the model in **c** by further learning the game-specific η^{s}_{other} 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 behaviours, with 100% completeness matching the predictive accuracy of the MLP model. All reported results were based on 10-fold cross-validation (see Supplementary Table 2 for details). Our focus was on the heterogeneity across games rather than the heterogeneity across participants.

decision-making domains, such as choice between risky lotteries[16,17], suggests that the parameters of behavioural models can be highly context sensitive, for example, because the complexity or cognitive difficulty of decisions varies across problems. Here, 'context' specifically refers to the configuration of the payoff matrix, meaning that people's behaviour may vary in response to different game set-ups. Each configuration serves as its own context, which allows context-dependent models to capture how strategic behaviour adapts to the particular game being played. The deep neural network model represents an extreme case of potential context dependence, as it is capable of forming its predictions based on the specific characteristics of each game. Using neural networks to capture context dependence has been successful in various domains of human behaviour[8,16,18–21]. Yet while our deep neural network 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 behavioural models with a neural network that is responsive to game-specific features. We focus on three key behavioural primitives: (1) the level of strategic sophistication (that is, *k*), (2) the player's level of noisiness (that is, η_{self}; neural QR model) and (3) the player's beliefs regarding the noisiness of others (that is, η^{s}_{other}; neural belief noise model). These behavioural primitives, especially when combined together, form the basis for a wide spectrum of behavioural models in the literature[1,4,12,13]. 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* (Supplementary Information section B).

As an example, consider the level-*k* neural QR model (Fig. 2c), in which the player's noisiness was predicted by a neural network. In this approach, an MLP was used to model the function $\eta_{self} = f_{MLP}$(game matrix), allowing the η_{self} parameter to vary across games. Once trained, the MLP can predict η_{self} for any given game matrix. By combining the strategic sophistication level *k* with the MLP-predicted η_{self}, the

https://doi.org/10.1038/s41562-025-02230-5

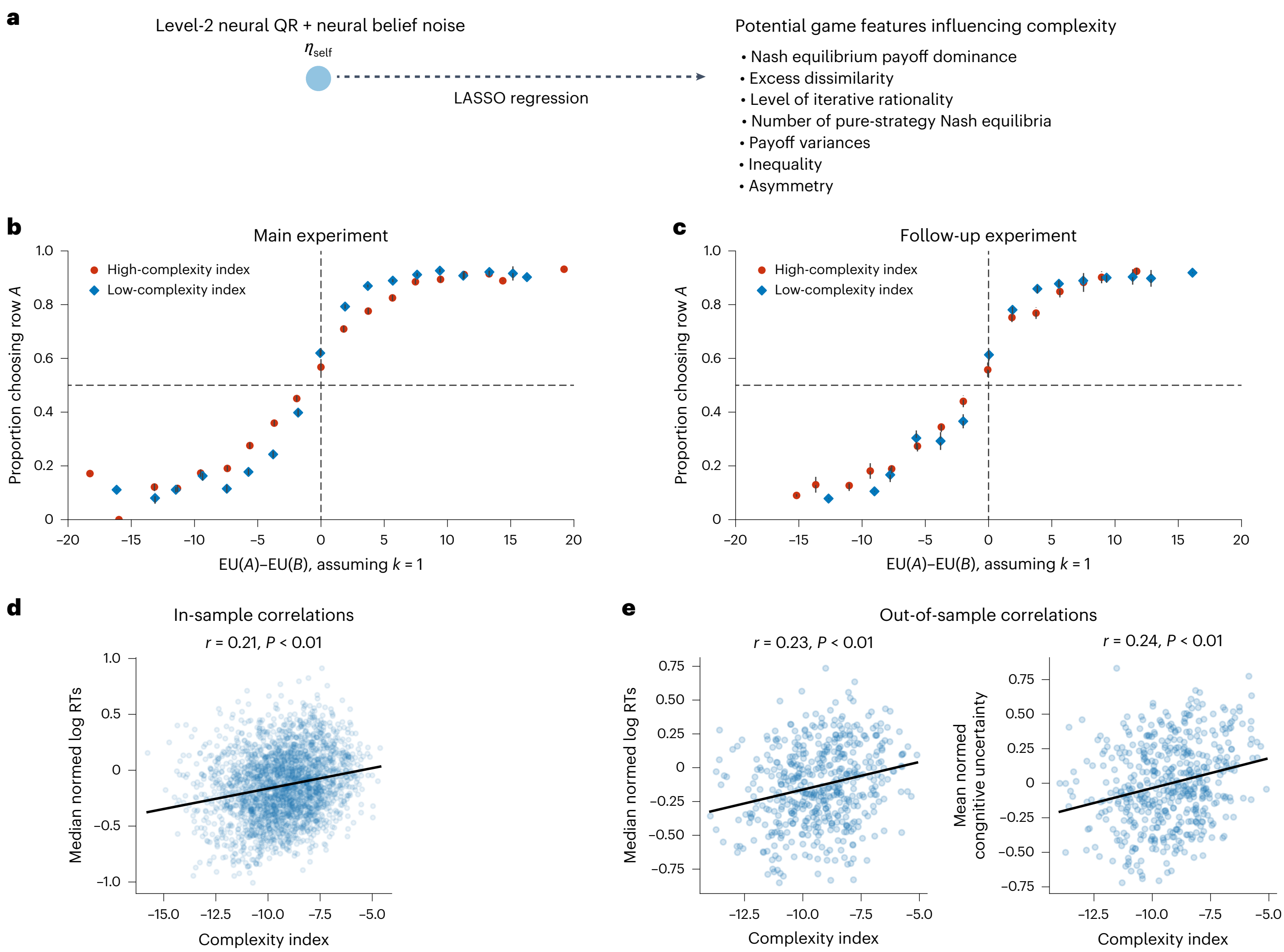

Fig. 3 | Developing an interpretable complexity index for strategic games. **a**, To construct an index of game complexity, we use LASSO regressions to learn game features that correlate with the game-specific η_{self} parameter that is estimated by the MLP in the level-2 neural QR and neural belief 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-complexity (low-complexity) games, determined by a median split on the complexity index. Error bars represent the s.e.m. **c**, The same psychometric function and effect of complexity was found in the follow-up experiment. Error bars represent the s.e.m. **d**, The complexity index shows a statistically significant correlation with RTs in the games of the main experiment. We used Pearson's correlation with two-sided significance tests to assess the linear relationship between game complexity index and RTs $r = 0.21$, $P < 0.01$. **e**, The complexity index generalized to the follow-up experiment and demonstrated statistically significant correlations with both RTs (Pearson's $r = 0.23$, $P < 0.01$) and cognitive uncertainty ratings (Pearson's $r = 0.24$, $P < 0.01$). Pearson's correlation was used with two-sided P values; no correction for multiple comparisons was applied.

level-*k* QR function iteratively applies the QR 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 behavioural models, the neural networks effectively learn a function that maps the game matrix to the parameters in the original behavioural model. Overall, the procedure of augmenting behavioural 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 Fig. 2e, integrating neural networks (that is, MLPs) into the level-*k* QR model to account for context dependence greatly 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 η_{self} and η^{s}_{other}. This context dependence (that is, 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 η_{self} or η^{s}_{other} across different games (Supplementary Table 2), we found that variations or context dependence in η_{self} play a more important role. Specifically, η_{self} has a mean value of 9.28 (95% confidence interval 9.14–9.42), whereas η^{s}_{other} has a mean value of 1.09 (95% confidence interval 1.03–1.15). 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, Fig. 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 QR model. We split the sample by the median estimated noisiness (η_{self}). Here, the noisiness was directly estimated using the complexity index (hence the interchangeable use of 'complexity index' and 'noisiness' in Fig. 3b), which we will develop later. 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 QR 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' behaviours are contingent on the specific game being played. We attribute this sensitivity to game complexity, by which we mean the cognitive difficulty of (1) forming beliefs about other players' strategies and (2) 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 analysed the predicted η_{self} values from the context-dependent level-2 QR model (Fig. 3a, the level-2 neural QR + neural belief noise model) and conducted a least absolute shrinkage and selection operator (LASSO) regression on a large set of structural game features (Supplementary Table 4). 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 (1) a measure of the cognitive difficulty of navigating trade-offs across different strategies, akin to work on dissimilarity and trade-off complexity in the literature on lottery choice[17,22]; (2) the variance and scale of payouts[16]; and (3) the inequality and asymmetry in payouts between players[23,24].

We collapsed these interpretable game features into a composite index of complexity (Supplementary Information section C). This index can be structurally interpreted as capturing the magnitude of the negated η_{self} predicted by game features. Because this index is defined on the basis of 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 showed a positive correlation with response times (RTs; Pearson's $r = 0.21$, $P < 0.01$, Fig. 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[25]. The results confirmed that the index robustly predicts out-of-sample behavioural outcomes. As shown in Fig. 3e, we replicated the positive correlation between RTs and the complexity index (Pearson's $r = 0.23$, $P < 0.01$). In addition, we observed a positive correlation between cognitive uncertainty and the complexity index (Pearson's $r = 0.24$, $P < 0.01$), 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 behavioural 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.

Discussion

Large-scale experiments and machine learning techniques have greatly facilitated our exploration of the vast space of cognitive mechanisms underlying strategic choices. Our findings indicate that both a player's own noisiness in responding to an opponent's behaviour and their beliefs about the opponent's noisiness are critical in determining deviations from the rational Nash equilibrium, RTs and uncertainty in strategic decisions. Moreover, allowing the degrees of noisiness to vary across different games (that is, context-dependent model parameters) improves the model's predictive accuracy. 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 behaviour in matrix games with differing levels of complexity.

A caveat to our work is that real-life strategic interactions are often less easily quantified than those analysed in the laboratory, making game matrices an incomplete representation of the full complexity of human strategic decision-making. In real-world scenarios, outcomes, opponents' intentions and strategies can all be ambiguous. Neural network models, particularly those capable of processing visual or natural language data, offer promising avenues for modelling strategic behaviour in more complex, real-life situations that may deviate from the structure of matrix games.

Overall, the key psychological insight from our study is that human strategic behaviours—including strategic choices, the time spent and the uncertainty of making those choices—can be effectively captured by a noisy best-response model. In particular, people's ability to identify a best response is heavily influenced by the complexity of the game and that this complexity is quantifiable. The key factors determining the complexity of the game, and thus the noisiness, include the efficiency feature (that is, Nash equilibrium payoff dominance) and the cognitive difficulty involved in identifying the best-response strategy (that is, excess dissimilarity and levels of iterative rationality). 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 behavioural game theory.

Methods

Main experiment

Game generation algorithm. The games2p2k dataset comprises 2,416 instances of 2 × 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 eight-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 1–50. Consequently, the set of all possible 2 × 2 games encompasses 50^8 games before considering permutations of game matrices.

To generate games from this space, we used a random generation process to create eight-item payoff matrices, where each item represents a random draw from a two-tiered uniform distribution $U[1, u]$ where $u \sim U[1, 50]$ (the tilde indicates that u is drawn from the uniform distribution). Next, we excluded games that lacked pure-strategy Nash equilibria and categorized each remaining game using Robinson and Goforth's topology for 2 × 2 games[11]. 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 (that is, 1,208 × 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-min experiment. These participants ranged in age from 18 to 86 years, with a median age of 37. Participants were required to be from the USA 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 US$2.00 (hourly rate of US$12.00), with the possibility of an additional bonus up to US$0.50 contingent on the result of a randomly selected game (1 point equals US$0.01). The experimental sessions were carried out in December 2023. Informed consent was obtained from all participants (Princeton University institutional review board (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 preexperiment multiple-choice quiz. After 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 played a game only 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 after 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 four 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 used 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-min experiment. Participant ages ranged from 18 to 88 years, with a median age of 35. The same filter used in the main experiment was applied here: participants were required to be from the USA 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 US$2.00 (hourly rate of US$12.00), with the possibility of earning a bonus of up to US$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 hidden until the participant interacted with it by clicking on the slider bar.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The datasets generated and/or analysed are available at https://osf.io/xrvaw.

Code availability

The code used to generate the results is available at https://osf.io/xrvaw.

References

1. McKelvey, R. D. & Palfrey, T. R. Quantal response equilibria for normal form games. *Games Econ. Behav.* **10**, 6–38 (1995).
2. Camerer, C. F. *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton University Press, 2011).
3. Gächter, S. Behavioral game theory. In *Blackwell Handbook of Judgment and Decision Making* (eds Koehler D. J. & Harvey, N.) 485–503 (Blackwell Publishing Ltd., 2004).
4. Weizsäcker, G. Ignoring the rationality of others: evidence from experimental normal-form games. *Games Econ. Behav.* **44**, 145–171 (2003).
5. McKelvey, R. D. & Palfrey, T. R. An experimental study of the centipede game. *Econometrica* **60**, 803–836 (1992).
6. Eyster, E. Errors in strategic reasoning. In *Handbook of Behavioral Economics: Applications and Foundations* (eds Bernheim, B. et al.) Vol. 2, 187–259 (North Holland, 2019).
7. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Quart. J. Econ.* **114**, 817–868 (1999)
8. Fudenberg, D. & Liang, A. Predicting and understanding initial play. *Am. Econ. Rev.* **109**, 4112–4141 (2019).
9. Golman, R., Bhatia, S. & Kane, P. B. The dual accumulator model of strategic deliberation and decision making. *Psychol. Rev.* **127**, 477–504 (2020).
10. Wright, J. R. & Leyton-Brown, K. Predicting human behavior in unrepeated, simultaneous-move games. *Games Econ. Behav.* **106**, 16–37 (2017).
11. Robinson, D. & Goforth, D. *The Topology of the 2×2 Games: A New Periodic Table* (Routledge, 2005).
12. Camerer, C. F., Ho, T.-H. & Chong, J.-K. A cognitive hierarchy model of games. *Quart. J. Econ.* **119**, 861–898 (2004).
13. Crawford, V. P., Costa-Gomes, M. A. & Iriberri, N. Structural models of nonequilibrium strategic thinking: theory, evidence, and applications. *J. Econ. Lit.* **51**, 5–62 (2013).
14. Murnighan, J. K., Roth, A. E. & Schoumaker, F. Risk aversion in bargaining: an experimental study. *J. Risk Uncertainty* **1**, 101–124 (1988).
15. Fudenberg, D., Kleinberg, J., Liang, A. & Mullainathan, S. Measuring the completeness of economic models. *J. Polit. Econ.* **130**, 956–990 (2022).

16. Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D. & Griffiths, T. L. Using large-scale experiments and machine learning to discover theories of human decision-making. *Science* **372**, 1209–1214 (2021).
17. Enke, B. & Shubatt, C. Quantifying lottery choice complexity. Preprint at https://doi.org/10.3386/w31677 (2023).
18. Dezfouli, A. et al. Disentangled behavioural representations. In *Advances in Neural Information Processing Systems* (eds Wallach, H. et al.) (2019).
19. Agrawal, M., Peterson, J. C. & Griffiths, T. L. Scaling up psychology via scientific regret minimization. *Proc. Natl Acad. Sci. USA* **117**, 8825–8835 (2020).
20. Hartford, J., Wright, J. R. & Leyton-Brown, K. Deep learning for predicting human strategic behavior. In *Advances in Neural Information Processing* (eds Lee, D. et al) (2016).
21. Marchiori, D. & Warglien, M. Predicting human interactive learning by regret-driven neural networks. *Science* **319**, 1111–1113 (2008).
22. Shubatt, C. & Yang, J. Similarity and comparison complexity. Preprint at https://arxiv.org/abs/2401.17578 (2024).
23. Chmura, T., Kube, S., Pitz, T. & Puppe, C. Testing (beliefs about) social preferences: evidence from an experimental coordination game. *Econ. Lett.* **88**, 214–220 (2005).
24. Bland, J. & Nikiforakis, N. Coordination with third-party externalities. *Eur. Econ. Rev.* **80**, 1–15 (2015).
25. Enke, B. & Graeber, T. Cognitive uncertainty. *Quart. J. Econ.* **138**, 2021–2067 (2023).
26. Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).

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

Author contributions
J.-Q.Z.: conceptualization, methodology, investigation, formal analysis, data curation, software, visualization, writing (original draft), and writing (reviewing and editing); J.C.P.: methodology, formal analysis, software, visualization, and writing (reviewing and editing); B.E.: conceptualization, methodology, formal analysis, supervision, validation, and writing (reviewing and editing); T.L.G.: conceptualization, methodology, formal analysis, supervision, validation, writing (reviewing and editing), resources, and funding acquisition.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-025-02230-5.

Correspondence and requests for materials should be addressed to Jian-Qiao Zhu.

Peer review information *Nature Human Behaviour* thanks James (R) Bland and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Corresponding author(s): Jian-Qiao Zhu

Last updated by author(s): Apr 13, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Experimental tasks were programmed using Dallinger 9.11.0.
Data analysis	Data analyses were conducted in Python 3.11.5 using the following packages: JAX 0.4.24, NumPy 1.24.3, Optax 0.1.9, pandas 1.5.3, and SciPy 1.11.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The datasets generated and/or analyzed are available at the following link: https://osf.io/xrvaw

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	In the main experiment, we recruited 4,900 participants via the Prolific Academic platform. Of these, 4,673 participants successfully completed the 10-minute task, including 1,942 males, 1,782 females, and 949 individuals who chose not to disclose their gender. In a follow-up experiment, an additional 1,013 participants were recruited through the same platform, with 1,008 completing the task (346 males, 416 females, and 246 who did not disclose their gender). Gender was self-reported via Prolific and was not linked to individual-level behavioral data. Accordingly, no sex- or gender-based analyses were performed.
Reporting on race, ethnicity, or other socially relevant groupings	Race, ethnicity, and other socially relevant groupings were not analyzed in this study.
Population characteristics	Participants in the main experiment ranged in age from 18 to 86 years (median age = 37), while those in the follow-up experiment ranged from 18 to 88 years (median age = 35).
Recruitment	All participants were recruited via the Prolific Academic platform. Eligibility criteria required participants to be located in the United States and to have completed at least 100 prior studies on the platform with a minimum approval rate of 95%.
Ethics oversight	Princeton University IRB number 10859: "Computational Cognitive Science" and Harvard University IRB number 16-1753

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Participants were instructed to play a series of economic games and to report their strategic choices, response times, and confidence ratings for each choice.
Research sample	Representative samples from the Prolific Academic platform.
Sampling strategy	Each participant completed a sequence of 20 games, randomly drawn from a pool of procedurally generated games.
Data collection	Behavioral data were collected online through participants' web browsers.
Timing	The main experiment was conducted in December 2023, while the follow-up experiment was conducted in April 2024.
Data exclusions	Participants with incomplete data were excluded from all subsequent analyses. Exclusion criterion for the follow-up experiment was preregistered at https://osf.io/xrvaw/.
Non-participation	A total of 227 participants dropped out or failed to complete the main experiment, while 5 participants did not complete the follow-up experiment.
Randomization	Participants were not assigned to experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.