nature human behaviour

ARTICLES

https://doi.org/10.1038/s41562-021-01207-4

Check for updates

Latent motives guide structure learning during adaptive social choice

Jeroen M. van Baar[1,2], Matthew R. Nassar[3,4], Wenning Deng[1] and Oriel FeldmanHall[1,4] ✉

Predicting the behaviour of others is an essential part of social cognition. Despite its ubiquity, social prediction poses a poorly understood generalization problem: we cannot assume that others will repeat past behaviour in new settings or that their future actions are entirely unrelated to the past. We demonstrate that humans solve this challenge using a structure learning mechanism that uncovers other people's latent, unobservable motives, such as greed and risk aversion. In four studies, participants (*N* = 501) predicted other players' decisions across four economic games, each with different social tensions (for example, Prisoner's Dilemma and Stag Hunt). Participants achieved accurate social prediction by learning the stable motivational structure underlying a player's changing actions across games. This motive-based abstraction enabled participants to attend to information diagnostic of the player's next move and disregard irrelevant contextual cues. Participants who successfully learned another's motives were more strategic in a subsequent competitive interaction with that player in entirely new contexts, reflecting that social structure learning supports adaptive social behaviour.

Humans spend a substantial amount of time trying to predict how others will behave. Which millennial has not agonized over the perfect emoji to round off a text message, hoping to elicit the desired response from a love interest? In a professional context, figuring out how a colleague will treat another co-worker during a disagreement can clarify whether it is worthwhile to entertain future collaborations with that colleague. Even in large-scale societal coordination problems such as political activism[1], climate change mitigation[2] and disease control[3], reliable knowledge about the future behaviours of others can be critical to reaching the desired outcome.

Despite the ubiquity of social predictions in daily life, little is known about how humans solve this difficult task. Consider trying to predict whether a co-worker will be a good collaborator on a group project. We cannot assume that our co-worker will act just as they did in the past, because each project is different (the people involved, the current economic conditions and so on). Instead, we must selectively generalize what we know about our co-worker to this new situation. If we generalize too little from past experience, we deny ourselves potentially valuable social information, but generalizing too much makes us slow to respond to the changing behaviour of others. To solve this generalization dilemma during social prediction, humans probably build parsimonious mental models of others' behaviour that are easy to update yet provide relatively accurate predictions in diverse social settings[4]. How do we construct and use these mental models to predict the behaviours of others?

Here, we examine the possibility that motives—a term we use to encompass a range of stable preferences and personality traits spanning both moral and non-moral concerns—form the building blocks of these mental models of others' behaviour. Motives, such as greed[5], inequity aversion[6,7] and risk aversion[8,9], are plausible components of such models because they can predict behaviour across contexts[10], differ reliably between people[10–14] and are inferred from an early age[15–17]. Studies of action understanding suggest that people can infer another's motives by observing their actions[18]. For example, another's food preference can be estimated by using Bayesian inference to compute which of several alternative foods is most likely to be selected[19]. However, this process becomes particularly challenging in naturalistic social settings because one behaviour can arise from several motives[12,20]. For instance, individuals may co-operate because they care about fairness[6], or because they want to appear co-operative to others (reputation management)[21], or simply because they feel good when helping[22]. To solve this inference problem, people must disambiguate the latent structure of another's choice patterns—that is, their unobservable overarching goals and motives—across diverse social contexts.

Because it is unknown which cognitive mechanism supports learning of others' latent motives, we test computational accounts of this structure learning process, which can identify mechanisms that are not readily seen by examining behaviour alone. Specifically, we evaluate whether a feature-based reinforcement learning system, thought to account for non-social, low-dimensional structure learning[23–29], can also support structural learning during social prediction. When applied to social cognition, feature-based reinforcement learning would require participants to learn latent, unobservable features that can account for a person's behaviour, such as a motive (for example, greed or risk aversion). If successful, latent structure learning would allow for greater efficiency in social prediction by selectively attending[23,25,30,31] to task information that is specifically relevant to the inferred motive. To assess whether this type of model applies to social prediction, we combine both behavioural and eye-tracking experiments with computational modelling of feature-based reinforcement learning.

In our experimental Social Prediction Game, participants are tasked with predicting the choices of another Player interacting with anonymous Opponents (Fig. 1a) in four types of economic games, each characterized by distinct tensions between potential gains and losses: the Harmony Game, the Snowdrift Game, Stag Hunt and Prisoner's Dilemma (Fig. 1b; example payoff matrices for each game type are shown in Supplementary Fig. 1). As Players move between

[1]Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, USA. [2]Trimbos Institute, Netherlands Institute of Mental Health and Addiction, Utrecht, the Netherlands. [3]Department of Neuroscience, Brown University, Providence, RI, USA. [4]Carney Institute for Brain Science, Brown University, Providence, RI, USA. ✉e-mail: Oriel.FeldmanHall@Brown.edu

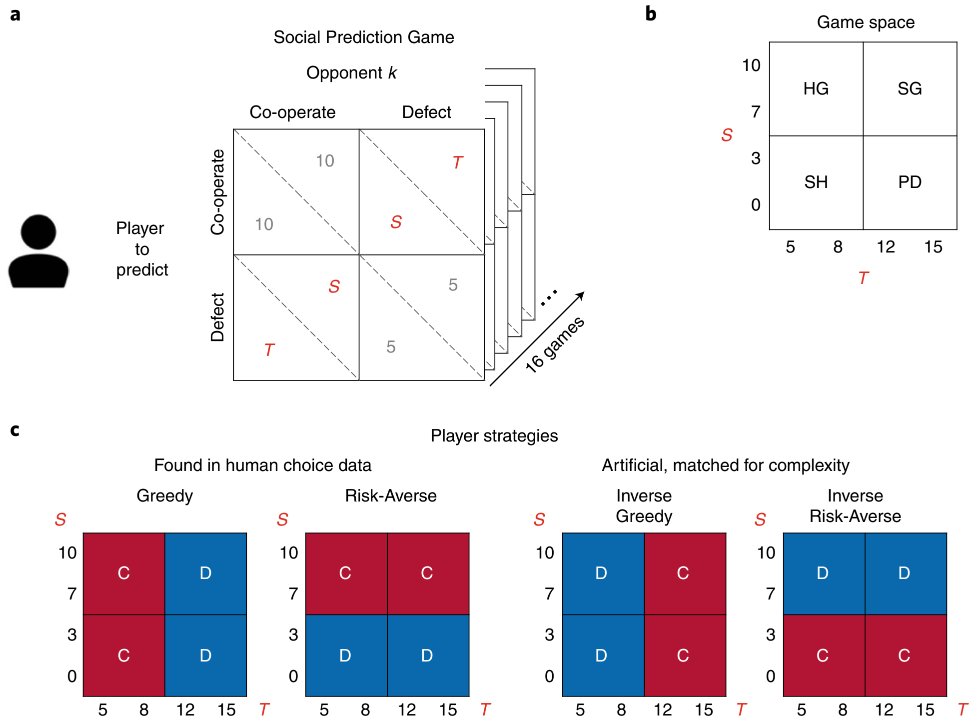

Fig. 1 | The Social Prediction Game. a, In each trial, the Social Prediction Game is presented as a symmetric payoff matrix describing the single-shot interaction between a Player (who is tracked across 16 consecutive trials) and anonymous Opponents (who are different in each trial). For each of the four possible outcomes (co-operate–co-operate, co-operate–defect and so on), the Player earns the value in the lower triangle, and the Opponent earns the value in the upper triangle. **b**, The values of S and T vary from trial to trial and are drawn from a 4×4 game space such that each trial belongs to one of four canonical economic game types: Harmony Game (HG), Snowdrift Game (SG), Stag Hunt (SH) and Prisoner's Dilemma (PD). In the HG, S is high and T is low, meaning that all payoff-maximizing Players have a strong incentive to co-operate regardless of their social preferences or risk attitudes. In the SG, by contrast, T is higher than 10, meaning that the greatest payoff can be won through unilateral defection. This motivates Greedy Players to defect in the SG, while Risk-Averse Players prefer the relatively high S over the mutual defection payoff of 5. In the SH, the greatest possible payoff is 10, motivating Greedy Players to co-operate, while Risk-Averse Players avoid the low S by defecting. In the PD, defection always yields the greatest payoff for the individual. Refer to Supplementary Fig. 1 for example payoff matrices of each game type. **c**, Across these game types, the Player's true behaviour (co-operate (C) or defect (D)) depends on the Player's underlying strategy, which is deterministically programmed in each block of 16 trials.

games, the payoffs associated with co-operating and defecting shift. This causes the tensions associated with different underlying motives to vary from game to game, revealing several distinctly structured patterns of behaviour[13]. For example, if a Player is motivated by risk aversion, they will choose to co-operate in a Snowdrift Game (where both players are tempted to defect, but mutual defection yields the lowest payoff) but defect in a Stag Hunt (where both players are encouraged to co-operate, but unilateral co-operation yields the lowest payoff). Although these decisions seem contradictory at first blush, they are consistent at a latent motive level, as these actions are both risk-averse in that they forgo the highest possible payoff in favour of minimizing potential losses[13,32].

Other motives, such as greed or envy, yield differently structured choice patterns across games[13] (Methods). This has two implications. First, when moving from game to game, the same action (for example, co-operate) can result from different motives, which means that a subject cannot rely only on actions to infer motives but instead must consider a Player's action relative to the incentives of the current game. Second, the same motive (for example, risk aversion) can produce different actions depending on the social tensions in a given game. Together, these features of our task mirror the generalization dilemma of social prediction. The hidden structure of Players' choices in our paradigm thus allows us to test whether and how participants selectively generalize knowledge about others' behaviour to new social settings. We hypothesized that (1) people generalize information about others' decisions between economic games on the basis of latent motives, (2) generalization is implemented by a feature-based reinforcement learning system that supports an increasing focusing of attention on information relevant to the inferred motives and (3) people use these inferred motives to make more adaptive choices in different competitive contexts.

Results

In study 1, 150 participants played the Social Prediction Game, where they observed four different Players playing 16 trials of four different economic games. In each trial, the participants were asked what the Player would choose to do (co-operate or defect) in the current game, and to rate their confidence in their prediction (see the task screenshot in Supplementary Fig. 2). The games between the Players and their Opponents were presented as a 2×2 payoff matrix (Fig. 1a), where two of the four possible outcomes yield the same payoffs across all trials: (10,10) for mutual co-operation and (5,5) for mutual defection. The payoffs assigned to the other potential outcomes, labelled S for sucker's payoff and T for temptation to defect, are drawn from a 4×4 grid (the game space) such that each of the 16 trials is a new game with a unique (S,T) combination. The varying (S,T) payoff values yield four canonical economic game types, which are represented in four quadrants of the (S,T) game space (Fig. 1b and Methods). For example, the classic Prisoner's Dilemma is found at $S<5$ and $T>10$, where all players are best off defecting regardless of their opponent's choice[33]. Due to these regularities in game incentives, human choice data reveal a limited number of distinct patterns of correlated choice (that is, strategies) across the 16 games, with each strategy optimizing for a distinct human motive, such as risk aversion or envy[13].

We programmed the Players to behave in line with one of four deterministic strategies: two found in human choice data at similar rates (namely, a Greedy strategy that maximizes the maximal payoff and a Risk-Averse strategy that maximizes the minimal payoff; respectively 20% and 21% of the population[13]) and two not found in human choice data but matched for statistical complexity (that is, Inverse Greedy and Inverse Risk-Averse; Fig. 1c). These motives were selected on two conditions. First, greedy and risk-averse motives are observed at comparable rates (~20%) in the general population[13], which buffers against the possibility that a more dominant motive—such as envy—might unduly bias learning in the Social Prediction Game relative to a less dominant motive. Second, because greed and risk aversion have equally complex patterns of decisions across the game space (Fig. 1c), we can draw inferences about social learning processes without worrying about confounds relating to statistical complexity.

Selective generalization of social information. How might participants approach the Social Prediction Game? One possibility is that they expect a Player to simply repeat their past behaviour due to stable preferences for co-operation[10]. This could be thought of as basic reinforcement learning, where the participant learns the value of predicting 'co-operate' and 'defect' for the current Player without distinguishing between different games (an approach doomed to fail in the more complex Social Prediction Game). Another possibility is that participants refrain from generalizing across games at all, because each trial is unique. Since all Players co-operate and defect on half the trials, both these strategies would yield on average 50% accuracy in our task. However, the observed accuracy was significantly greater (59.1% ± 9.1% (s.d.); two-tailed one-sample *t*-test: $t(149) = 12.2$, $P < 0.001$, Cohen's $d = 1.00$).

A third possible strategy is naïve statistical learning, whereby participants detect the mapping between *S* or *T* and the Player's choices (for example, learning that Inverse Risk-Averse co-operates when $S < 5$). Such a strategy reflects how participants learn latent structure in non-social tasks containing abstract stimuli such as coloured shapes and fractals[26,27]. If true, task performance should be equal across all Player strategies, as each strategy is a step function with a single change point on the *S* or *T* dimension (Fig. 1c). However, performance was much higher for human than artificial strategies (Greedy and Risk-Averse: average accuracy, 71.6% ± 10.5%; Inverse strategies: 46.5% ± 12.4%; two-tailed paired-samples *t*-test: $t(149) = 22.0$, $P < 0.001$, $d = 1.80$; Fig. 2a). This reveals that naïve statistical learning does not capture learning in this task, and suggests that how we learn about other people differs from how we learn about objects (such as coloured shapes and fractals) in the non-social world.

These findings demonstrate that (1) the participants selectively generalize information across games by learning a latent structure that links game incentives to choice, and (2) the participants hold strong prior expectations about the nature of this latent structure, as they were able to predict the choices of the human Players but not artificial ones. We next used computational modelling to illuminate how the participants acquired and applied this latent structure when generalizing information about others' decisions across games.

Motives guide social structure learning. To determine how participants generalize information across games during social prediction, we built a formal model of structural learning that could capture how the participants map the *S* and *T* game variables onto choice. We leveraged recent developments in feature-based reinforcement learning, where agents do not learn the value of stimuli per se (for example, apples or oranges) but rather the decision relevance of stimulus features (for example, their size or colour)[23,28]. We adapted this class of model by allowing it to learn not just over observable task features *S* and *T* but also over unobservable features such as game type or a Player's motives. We then evaluated various sets of features to test how participants generalize learning across trials.

Task behaviour was best captured by a feature set that generalized information across games according to four human motives whose role in social decision-making is well-documented: Co-operativeness[34], Greed[5], Risk Aversion[9,35] and Expected Value (EV) maximization[36,37] (motive-based structure model; Fig. 3a and Methods). Model comparison using the Bayesian information criterion (BIC)[38] showed that the average fit of this motive-based model was superior to several alternative models, which together systematically covered a range of regular behaviour patterns across the (*S*,*T*) game space. First, corroborating our earlier observation that participants generalize learning across trials, our motive-based model outperformed a model that treated each trial as unique and thus learned over 16 game features and an intercept (no structure; $\Delta\text{BIC} = -55.41 \pm 25.2$ (mean ± s.d.), two-tailed Wilcoxon sign-rank test, $W = 7$, $P < 0.001$, $d = -2.20$). Second, our motive-based model outperformed a plausible alternative model: one that recognizes how Player choices are stable within each game type (Harmony Game, Stag Hunt and so on) and thus learns over four game features and an intercept, generalizing learning within (but not selectively across) game types (game type-based structure; $\Delta\text{BIC} = -5.95 \pm 10.23$, $W = 1641$, $P < 0.001$, $d = -0.58$; Fig. 3b). Moreover, posterior predictive checks showed that the motive-based structure model qualitatively reproduced key features of participant data (Figs. 2a and 3c). These results suggest that the participants generalized learning across game types, with motives as a likely basis for such generalization.

To further test whether the participants used the four selected canonical motives to selectively generalize learning across games, we compared our motive feature set with 5,000 additional feature sets consisting of randomly generated pseudo-motives (achieved by shuffling the four canonical motive features as defined across the 16 (*S*,*T*) 'tiles' in the game space). This approach can be thought of as densely sampling the space of possible bases for generalization, which allowed us to test for evidence of motives we had not considered ourselves. Although the pseudo-motives could all selectively generalize learning across games—the hallmark feature of our motive-based model—no combination of them accounted for the participant data as well as the motives-based structure model, which stipulated four canonical motives (Supplementary Results 1). In fact, the randomization procedure recovered our theoretically guided set of motives, suggesting that these motives reflect the mental structure that participants use when generalizing to new contexts. Together with a number of model quality checks (Supplementary Results 2–4), these results support our hypothesis that participants rely on a feature-based reinforcement learning mechanism that leverages latent motives as a mental scaffold to generalize what they know about another person across social settings.

Complex mental models allow accurate predictions. Our observation of structure learning in a social prediction task poses an interesting puzzle. On the one hand, it is likely that structure learning improves social prediction when facing a variety of latent motives in one's social environment[12,13]. On the other hand, identifying a greater number of latent motives requires updating the predictive value of several motives at once, which means paying attention to many different task features that each inform different motives, such as the values of *S* and *T*. This poses a trade-off in which participants may consider only a subset of the four motives in their mental models of others' behaviour. To test this, we again fit the motives model to the participants' prediction data, this time comparing different model versions that each contained a unique subset of motives (that is,

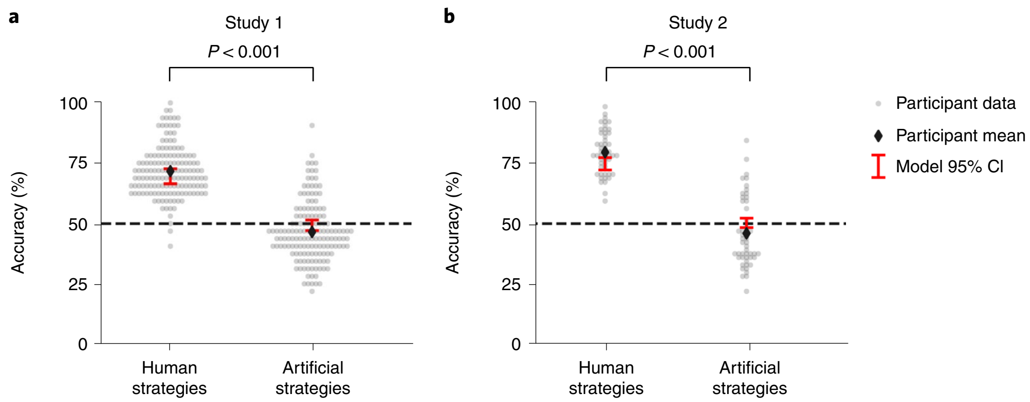

Fig. 2 | Behavioural results. a,**b**, In studies 1 (**a**) and 2 (**b**), task performance for inferring human strategies (Greedy and Risk-Averse) was greater than for inferring artificial strategies (Inverse Greedy and Inverse Risk-Averse), and the computational model accurately reproduced this effect. Each grey dot represents a participant. The black diamonds indicate the average participant accuracy. The red error bars are the 95% confidence intervals (CIs) of average accuracy obtained from simulating the fitted model 1,000 times per participant.

only Co-operativeness, only Greed, Co-operativeness + Greed and so on). We then ran model comparisons at the subject level, again using the BIC as a criterion to reward model accuracy and penalize complexity (that is, the number of motives in a given model).

Confirming our hypothesis, we found that the predictions of all participants were best accounted for by a personalized model that specified only a subset of the four motives rather than by the full model containing all four motives (ΔBIC = −14.9 ± 10.5, W = 0, P < 0.001, d = −1.42; Fig. 3b). This was true even when comparing model error without penalizing complexity (difference in sum of squared model error (SSE), −0.33 ± 0.63; W = 2,291; P < 0.001; d = −0.53). The most common motive subset (that is, the set of Co-operativeness and Greed) had greater average model error when fit to the entire group of participants than the full model with all four motives, although this difference was not statistically significant (Co-operativeness + Greed versus All Motives: ΔSSE = 0.42 ± 1.57, W = 4,613, P = 0.085, d = 0.27). This suggests that there was individual variation in the motives that our participants considered in the Players.

The participants differed on the number of motives they considered: 30% of participants considered only a single motive, 50% considered two motives and 20% considered three motives simultaneously. Participants who built more complex mental models showed superior task performance but also spent more time making their predictions (Supplementary Results 5), revealing a speed–accuracy trade-off of structure learning.

Self-anchoring guides motive inference. The participants also differed in what motives they considered: 87% of participants considered Greed in their mental model, 51% considered Co-operativeness, 31% considered EV and only 21% considered Risk Aversion. Accordingly, task performance was much higher when predicting the Greedy Player (average accuracy, 87.0% ± 12.9%) than when predicting the Risk-Averse Player (average accuracy, 56.3% ± 19.4%; two-tailed paired-samples t-test: $t(149)$ = 14.7, P < 0.001, d = 3.17).

To test what drove these individual differences in motives considered, we compared participants' best motive subset with their own decision strategy if they were to have played the game themselves (which participants reported during the task instructions). The Co-operativeness motive was more commonly considered in others by participants who reported their own decision strategy to be co-operative during the task instructions (one-tailed binomial test, P = 0.010; Fig. 3d), and the same was true for the Risk Aversion motive in participants who had reported a risk-averse strategy (P < 0.001). Predicting the actions of a Player who shares the same strategy as oneself led to highly accurate predictions (that is, Greedy and Risk-Averse participants were on average 87.0% accurate on the corresponding task blocks). However, even though self-anchoring led to a boost in accuracy, the participants were still able to predict the actions of Players that had different strategies at 69.4% accuracy on average, which is significantly better than guessing ($t(262)$ = 14.0, P < 0.001, d = 0.86). These self-anchoring effects suggest that our own motives help shape our constrained mental models of others' behaviour, illuminating the role of informed priors in social structure learning.

Social structure learning facilitates attentional focus. Although participants with more extensive mental models were slower to make social predictions, this relative time cost decreased between early trials (1–8) and late trials (9–16) (Supplementary Results 5), reflecting the fact that structure learning involves gradually simplifying complex decision problems to speed learning[30]. Specifically, it has been suggested that structure learning facilitates an attentional focus on information relevant to the (learned) latent structure of the task[23,25]. We tested this mechanistic hypothesis in study 2, where 50 participants played the same game—behaviourally replicating the findings from study 1 (Supplementary Results 6)—only this time in the laboratory and while undergoing concurrent eye-tracking. We hypothesized that as participants learned the motives of the Player, they would focus their attention on the task features relevant to those specific motives, namely T for the Greedy Player and S for the Risk-Averse Player (Fig. 1c). Note that the payoff matrices look identical for all Player types. Therefore, any difference in attentional patterns for a given game between blocks must result from learning in previous trials in that block, providing evidence that structure learning guides attention.

Confirming our prediction, participants spent a greater percentage of the trial looking at T relative to S in the Greedy block than in the Risk-Averse block (two-tailed paired-samples t-test, $t(49)$ = 5.11, P < 0.001, d = 0.73; Fig. 4a). To test whether this gaze difference developed as a function of structure learning, we split the Greedy and Risk-Averse blocks into four mini-blocks of four consecutive trials in which each game type was visited once, and we ran a mixed-effects regression with random subject intercepts to test whether Player strategy (Greedy/Risk-Averse) interacted with mini-block number to predict relative gaze (S versus T). The strategy–mini-block interaction was significant ($F(1,347)$ = 16.9, P < 0.001), with the greatest shift in attention taking place in the first four to eight trials (Fig. 4b), confirming that the attentional shift between T and S was conditional on learning about the Players'

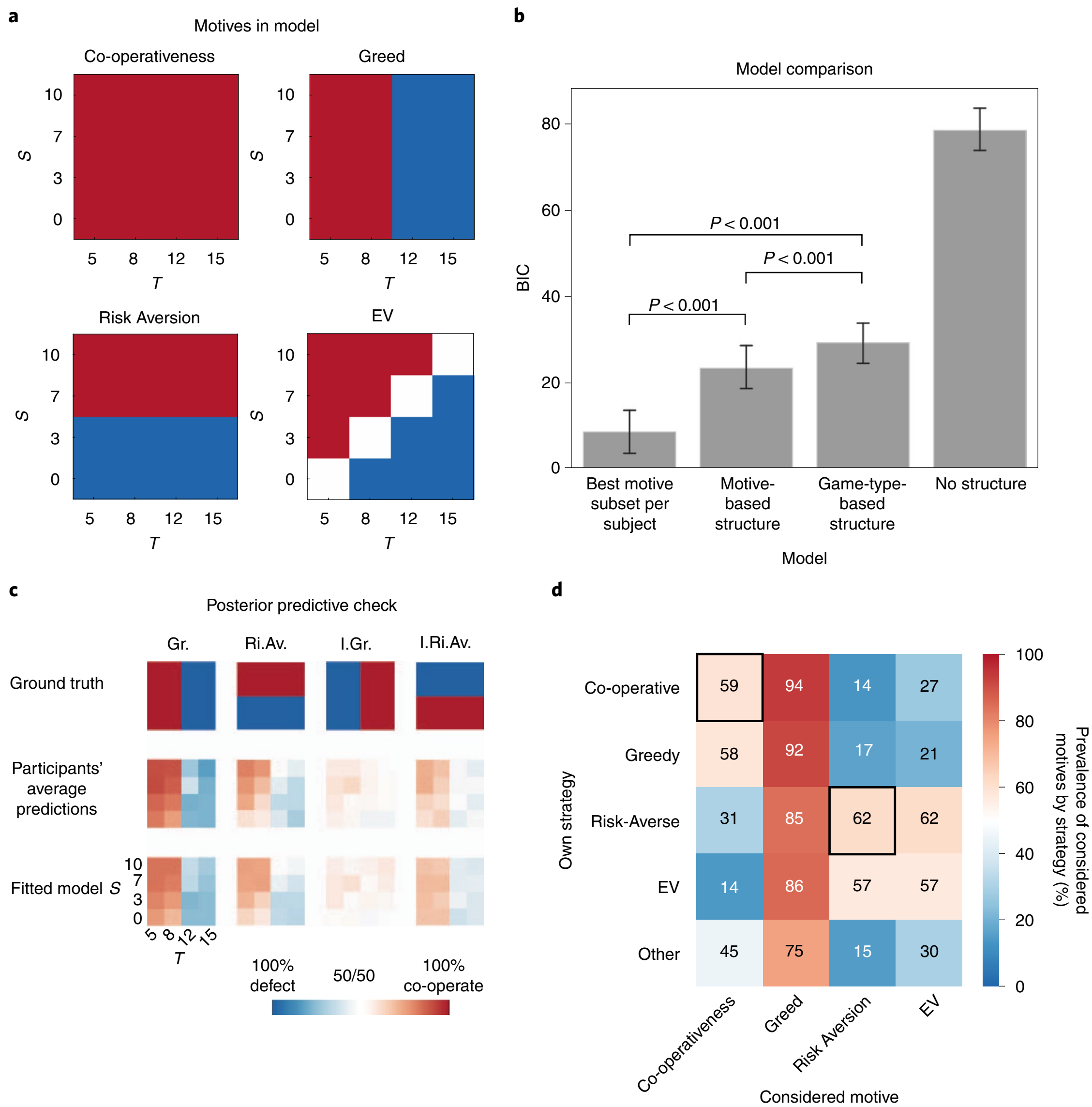

Fig. 3 | Computational models. a, The four canonical motives included in the winning model make different behavioural predictions across the (S,T) game space. **b**, Model comparison for study 1 showed that each participant's behaviour was best described by a model including a subset of the four canonical motives. The bars represent the average BIC over the sample for each model. The error bars represent bootstrapped 95% CIs. $P<0.001$ in two-tailed, non-parametric Wilcoxon sign-rank tests. **c**, Posterior predictive checks show that the model captures the fine-grained patterns of observed behaviour in study 1 as a function of Player strategy, S and T. Gr., Greedy; Ri.Av., Risk-Averse; I., Inverse. **d**, Self-anchoring in motive inference. This similarity matrix shows the relationship between participants' own strategies (rows) and the motives they consider in the Players (columns). While Greed was considered by almost all participants, Co-operativeness ($P=0.010$ in a binomial test) and Risk Aversion ($P<0.001$) were considered more often by participants who had this motive as their own decision strategy (key tiles highlighted by black outlines).

strategies. This effect was also found at the subject level: participants who considered Risk-Aversion in their mental model spent more time looking at information relevant to a Risk-Averse Player (S) than participants who did not consider this motive (Supplementary Results 7 and Fig. 4b).

Furthermore, our eye-tracking data provide some insight into two theoretical claims. First, the fact that participants who shifted their gaze more to the diagnostic information (T for Greedy, S for Risk-Averse) also performed better at the task (Risk-Averse block, correlation between relative gaze to S versus T and accuracy: $r(48)=0.37$, $P=0.008$; ceiling effect in Greedy block with on average 14 of 16 trials correct) suggests that structure-learning-based attentional focus may support adaptive decision-making[23,25]. However, we cannot rule out the possibility that increasingly appropriate attention is an epiphenomenon of making correct choices, and other measurements during the decision process (for example, mouse-tracking) would be needed to tease this apart. Second, the gaze data support the validity of the feature-based reinforcement learning framework for social structure learning, as our computational model was able to predict gaze over time as a function of the participant's beliefs about the Player's motives (Supplementary Results 7 and Fig. 4c, which qualitatively validates our model).

Inferred motives are robust to different Player types. Having established that a motive-based structure learning model can account for participants' predictions and eye movements in the Social Prediction Game, we aimed to test the generalizability of the model to different Player types. Specifically, since the participants

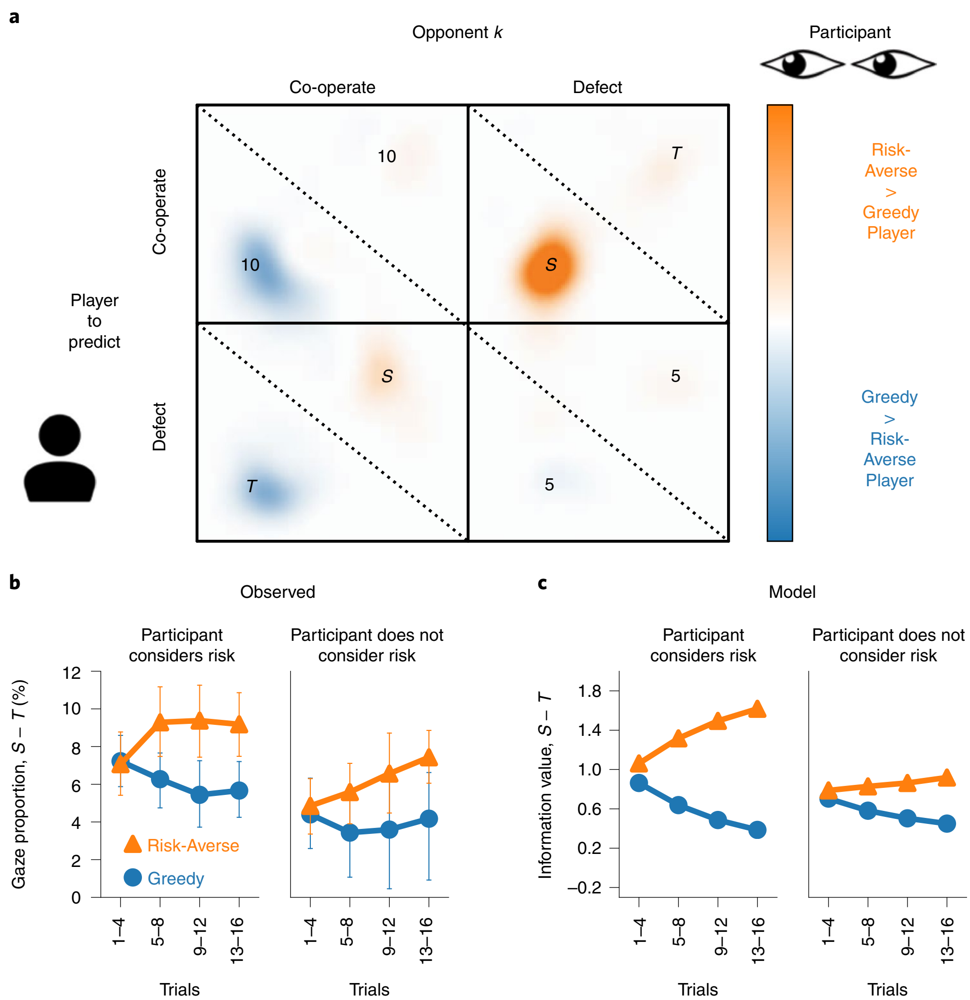

Fig. 4 | Eye-tracking results from study 2. a, Participants selectively looked at information diagnostic of a Player's strategy—that is, *T* for Greedy (blue) and *S* for Risk-Averse (orange). The heat map colours illustrate the direction of the effects and have no absolute meaning, which is why the colour scale is not quantitatively labelled. The statistics in the accompanying text quantify these effects. **b**, Selective attention developed over time as the participants learned the Player's motives. **c**, The computational model was fit only to choice and confidence data but qualitatively reproduced the participants' gaze shift over time as displayed in **b**.

were clearly best at predicting the Greedy Player in study 1 (Fig. 3c), we wanted to test whether high performance was also possible for other types of Players and whether our model could capture such performance. We therefore ran a third experiment ($N=148$) using the Social Prediction Game, this time presenting the participants with all four Player strategies found in human choice data[13], which we term Greedy, Risk-Averse, Co-operative (always co-operates) and Envious (co-operates only if it is guaranteed that one's own payoff will be equal to or greater than the Opponent's payoff). The results replicate our earlier observations while also validating our model (Supplementary Results 8). First, even though each of the 16 trials in a block was unique, the participants achieved above-chance prediction overall (average accuracy, $73.7\% \pm 8.8\%$ (s.d.); two-tailed one-sample t-test: $t(147)=32.8$, $P<0.001$, Cohen's $d=2.70$; Supplementary Fig. 3a). Second, performance was above chance for each individual Player type (Co-operative: $88.7\% \pm 13.2\%$, $t(147)=35.8$, $P<0.001$, $d=2.95$; Greedy: $85.0\% \pm 16.1\%$, $t(147)=26.4$, $P<0.001$, $d=2.18$; Risk-Averse: $54.1\% \pm 18.2\%$, $t(147)=2.73$, $P=0.007$, $d=0.22$; Envious: $67.2\% \pm 13.3\%$, $t(147)=15.8$, $P<0.001$, $d=1.30$; Supplementary Fig. 3b), revealing that accuracy is not limited to the Greedy Player.

Structure learning supports adaptive social choice. In studies 1–3, we found that people use a structure learning system to infer the latent motives of others in order to predict which social decisions those others will make. People probably engage in social structure learning (despite the associated effort) because it enables them to strategically adapt their behaviour to what others will do. However, deploying strategic social behaviour requires a deep and flexible understanding of the underlying latent structure of another's motives, which can enable a person to generalize another's motives to contexts that have never been experienced before.

To test this generalization effect in new situations, we carried out a fourth study in which the participants first played a block of the Social Prediction Game where the Player was programmed to be either Greedy or Risk-Averse. Afterwards, the participants played the Inspection Game[39,40] in the role of the Employer, and the Player took the role of the Employee. In the Inspection Game, the Employee receives a wage from the Employer and chooses to either 'work' (which creates revenue for the Employer but is costly for the Employee) or 'shirk'. Because the Employee receives the wage regardless of whether he or she actually works, shirking has the higher payoff if it goes unnoticed. However, the Employer can choose to pay a cost to inspect the Employee and withhold the wage if the

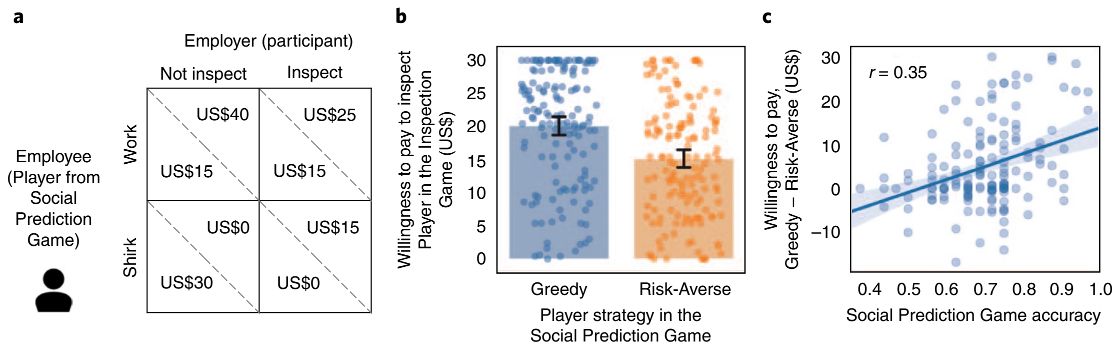

Fig. 5 | Task and results of Study 4. a, Payoff matrix of the Inspection Game, which was played after completing the Social Prediction Game. **b**, Despite an absence of feedback in the Inspection Game, the participants paid more to inspect Employees who exhibited Greedy strategies than those who exhibited Risk-Averse strategies in the preceding Social Prediction Game, revealing generalization of structure across tasks. Each dot is a participant; overlapping dots have darker colours. **c**, The difference in willingness to pay for inspecting between Greedy and Risk-Averse Players is predicted by learning performance in the Social Prediction Game.

Employee is found shirking. Given these rules, a Greedy Employee will always shirk to obtain the maximal payoff, while a Risk-Averse Employee will always work to avoid being caught shirking (Fig. 5a). Conversely, it is best for the Employer to pay for inspecting only if the Employee is likely to shirk. The natural mapping of motives in the Social Prediction Game to the Employee's behaviours in the Inspection Game create a strong test for whether successful structure learning in one context is adaptively used to behave more strategically in an entirely different context. Moreover, because participants play the Inspection Game with no feedback and are thus not aware of whether their partner is shirking or working, we can test whether participants are exclusively using their knowledge from the Social Prediction Game to make their decisions in the Inspection Game. Because the cost of inspecting changes from trial to trial (using a staircase design; Methods), we were able to identify a participant's indifference point for willingness to incur a cost to inspect. We predicted that successfully learning the motives of the Player during the Social Prediction Game should lead to paying more to inspect a Greedy Player than a Risk-Averse Player in the Inspection Game.

Confirming our hypothesis, the participants were willing to pay an average of US$4.99 more to inspect the Greedy Player (US$20.10 ± US$8.83) than to inspect the Risk-Averse Player (US$15.11 ± US$8.75; within-subject two-tailed Wilcoxon signed-rank test: $W = 2031.5$, $P < 0.001$, $d = 0.53$; Fig. 5b). This reveals that participants were successful in detecting another's latent motives in one context and generalizing this information to make more adaptive choices in new situations. Furthermore, a participant's willingness to pay was mediated by how well they inferred the motives in the Social Prediction Game. The better a participant performed in the Social Prediction Game, the more they paid to inspect Greedy relative to Risk-Averse Players ($r(151) = 0.35$, $P < 0.001$; Fig. 5c). These costly inspections paid off: the more participants distinguished between Greedy and Risk-Averse Employees in the Inspection Game, the more they earned as Employers (at a default inspection cost of US$15: $r(151) = 0.72$, $P < 0.001$). Finally, the more participants distinguished between Greedy and Risk-Averse in the Inspection Game, the more their self-reported inferences about the Players' strategies in the Social Prediction Game reflected the true underlying motives of Greed and Risk Aversion (Supplementary Results 9), validating both the Social Prediction Game as a social motive inference task and our conceptual labels of 'Greed' and 'Risk Aversion'.

Discussion

People routinely predict the behaviour of others across a dizzying array of social situations. We have shown that such sophisticated social prediction is achieved through structure learning: participants credit another person's social choices to co-operate or defect to latent, unobservable motives (such as greed and risk aversion) and use this motive-based abstraction to successfully predict decisions in entirely new interactions with different social tensions. Through structure learning, participants were able to disregard irrelevant contextual cues and attend to information that was diagnostic of the other player's future actions. Better use of structure learning led to more strategic behaviours in a subsequent competitive decision task with the same player, reflecting the adaptive value of accurate social prediction. Together, these findings provide a mechanistic and computational account of social prediction, illuminating how humans adaptively tackle a principal source of uncertainty in our social world[41]—other people.

This work establishes structure learning[23,24,28,30] as a critical mechanism for successful social cognition. In our experiments, popular learning models that are often relied on in non-social contexts (such as basic reinforcement learning over choice options and naïve statistical learning) could not account for the selective generalization of social information exhibited by the participants. Instead, the participants stripped down the complexities of the social exchange, leaving only a few latent dimensions by which to generalize what they had learned to entirely new contexts. These inferred dimensions differed across participants and were partly based on their own decision motives, which reveals a role for priors not found in non-social structure learning (for example, coloured shapes or fractals)[23,26,27]. Social structure learning was captured by a feature-based reinforcement learning model, originally developed for tracking relevant stimulus features in non-social decision tasks[23–25], which we adapted to learn over latent features of human social interactions such as unobservable motives. This establishes feature-based reinforcement learning as a potential algorithmic implementation of structure learning in social cognition, complementing existing accounts using non-parametric Bayesian models[19,42–44] and inverse reinforcement learning[45,46]. While posterior predictive checks from our modelling show that feature-based reinforcement learning provides a sufficient explanation of our data, we did not exhaustively explore the space of all possible models, and thus it is certainly possible that other algorithms (if equipped with appropriate motive representations) could do so as well. In line with the predictions of feature-based reinforcement learning[23], response times were slower in early trials, when the participants were considering multiple motives as potential causes of the Player's choices, and became faster as the participants homed in on a single motive and deployed selective attention to information relevant to that motive. This gradual narrowing of attention allowed the participants to limit cognitive resources use while still achieving accurate social prediction. In this way, motive-based structure learning resembles efficient processing in other cognitive domains, such as chunking in working memory[47,48]—a cognitive adaptation that leverages the inherent structure of the world.

This cognitive adaptation allows us to navigate our complex and evolving social environment by uncovering its latent structure, which is a better basis for generalization than simply attending to context cues or others' actions. Consider a toy example of a resistance hero who never lies to her parents but during enemy occupation staunchly denies being part of the opposition to protect her family. A superficial bookkeeping of this person's behaviour (she never lies) would fail to predict her actions in the novel context of war, while a deeper understanding of her latent motives (caring for her family) would facilitate accurate behavioural predictions. Similarly, in the Social Prediction Game, Greedy and Risk-Averse Players co-operated and defected equally often (50% of trials), and therefore generalizing only by their observable choices would yield identical (and highly uncertain) predictions in all new games. However, by abstracting and generalizing the Players' latent motives from the combination of game incentives and choices made, our participants correctly predicted divergent behaviour in an entirely new context, the Inspection Game[39,40]. After successfully learning latent structure, the participants paid for inspections only when it was valuable to do so, thereby earning significantly more money. Moreover, making more accurate predictions in the Social Prediction Game led to more strategic inspection (and more money earned) in the Inspection Game. This intimates that many strategic socio-economic behaviours that rely on predictions about others' future choices in novel contexts (such as competitive bargaining[49], market entry[50,51] and collective action[2,32,52,53]) are scaffolded by a motive-based structure learning mechanism. Given that social motives (for example, greed) vary between individuals in the population[10,12–14,54], effective inference of these motives in others is probably instrumental in achieving competitive and collaborative goals.

How do people construct and apply abstracted mental models of others' motives? Our data suggest that attention plays a key role in this process. Attention is a fundamental cognitive mechanism, as it affords optimal access to behaviourally relevant information with limited processing capacity[55]. Our findings show how structure learning guides attention and suggest that attention, in turn, supports accurate social prediction. In the Social Prediction Game, as in everyday social interactions, there were multiple cues that could be predictive of another's behaviour, from the Player payoffs S and T to the order of the games or even the initials of the Player. Structure learning allowed the participants to disregard superficial cues and attend to information relevant to the Players' latent motives. Conceptually, this process of attentional honing may facilitate accurate social prediction with limited effort if the inferred motives are correct, but incorrect structure learning may cause counterproductive attention to irrelevant information. In our experiment, participants who did not consider risk aversion also failed to shift their attention to S during the Risk-Averse block and instead kept looking at T, thereby missing out on information predictive of the Player's choices. This suggests that what we can learn about other people is limited by our expectations. Our mental models of others effectively constrain what we pay attention to, causing blind spots that prevent us from properly updating our beliefs and potentially contributing to stereotyping[56,57] and confirmation bias[58].

Our findings have implications for understanding these social behaviours in the real world. For example, police officers may disproportionally target racial minorities[59] because of a bias in information search, where inaccurate mental models of minorities' motives cause police officers to misinterpret behaviour or ignore exonerating information. Future work can directly test such hypotheses, using our tasks and model to disentangle the complementary roles of prior beliefs, information search and structure learning in social prediction. The current work thus provides new avenues for understanding failures of social prediction observed in everyday life.

Methods

Study 1 procedure. All studies in this paper were approved by the Brown University Institutional Review Board. In study 1, 150 participants (95 males, 52 females, 3 no response; mean age, 35.4 ± 10.0 yr) participated via Amazon Mechanical Turk (MTurk) in exchange for monetary compensation. The sample size was chosen such that key effects from smaller pilot studies could be observed with high statistical power. All participants gave written informed consent before starting the experiment. The task was written in Javascript and made accessible using Psiturk v.2.3.0 (ref. [60]). The participants first read the instructions and were quizzed to ensure their understanding and filter out potential bots. The participants were then asked to indicate for each game type in the Social Prediction Game how they themselves would choose, from which we estimated the participants' own decision strategies. They then completed the Social Prediction Game.

Task. The participants played four blocks of the Social Prediction Game, each block with a different Player, and were tasked with predicting the choices of this particular Player across 16 consecutive economic games. The Players always played single-shot against anonymous Opponents. Each game was presented as a 2×2 payoff matrix (Fig. 1a) where the Player and Opponent each have two choices: co-operation and defection. In the task, these choices were labelled by arbitrary colour words (such as blue or green) whose mapping to co-operation and defection changed on every task block.

The games varied on two features central to social interactions: risk of co-operating (here operationalized as S) and T (Fig. 1b). At $T < 10$ and $S > 5$, the games fall under a class of Harmony Game, where each player's payoff-maximizing action aligns with the jointly payoff-maximizing action, and thus no conflict arises except through potential envy[61]. At $T > 10$ and $S > 5$, the games are classified as Snowdrift Games (also known as Volunteer's Dilemmas), which are anti-coordination games where unilateral defection is preferable to mutual co-operation, but mutual defection yields the smallest payoff for all[62]. At $T > 10$ and $S < 5$ lie the Prisoner's Dilemma games, which are characterized by a high value of T even if one's opponent defects as well, and co-operation is risky as unilateral co-operation yields the lowest possible payoff[33]. At $T < 10$ and $S < 5$, the games are Stag Hunts, in which mutual co-operation yields the highest payoff for both, but co-operation is risky as unilateral co-operation is met with the lowest payoff[63].

The task of the participants was to indicate, in each trial, what they believed the Player would choose to do in the current game, and to rate their confidence in this prediction on an 11-point scale from 0% to 100% (10% increments). They received feedback on every trial indicating whether their prediction was correct or not, and earned a US\$0.01 bonus for every correct trial. At the end of 16 trials (one block), the participants self-reported what they believed the Player's strategy was using a free-response answer box. After four blocks, the total earned bonus was presented to the participants and added to the base payment. The participants were then taken to a survey hosted on Qualtrics to finish the experiment.

Model feature sets. We built three versions of our computational model with different types of features that each reflect a different psychological model about how participants might carve up the 4×4 game space to generalize information across trials.

Model 1 (motive-based structure) learns over four features representing psychological motives that span the entire game space. This reflects the hypothesis that participants realize that the Player's behaviour across all games is driven by overarching motives, which is used to generalize learning across games to improve predictions. Model 1 can support sophisticated generalization across the game space whereby co-operation in one game type predicts defection in another—a common pattern in human choices caused by the fact that consistently adhering to one motive can lead to different choices in different games[13,32]. As features in this model, we included four motives whose influence on behaviour in economic games is well-documented: Co-operativeness, Greed, Risk Aversion and EV (Fig. 3a). The Co-operativeness motive always chooses to co-operate and is based on common social norms that prescribe such behaviour[34]. Greed chooses the maximal payoff that can be obtained[5]. Risk Aversion chooses to maximize the minimal payoff[35] (optimizing the worst-case outcome). The EV motive follows a simplified and risk-neutral model of rational choice under uncertainty[36,64] that optimizes over the average of the two possible outcomes associated with a co-operate or defect decision, thus co-operating if $(10 + S) > (5 + T)$. The Greed, Risk Aversion and EV motives can all be understood as maximizing expected utility under different assumptions about the Opponent's choice: Greed maximizes expected utility if one believes that the Opponent will surely co-operate (sometimes termed optimism[13]), Risk Aversion maximizes expected utility assuming the Opponent will defect (pessimism) and EV maximizes expected utility under complete uncertainty (50% probability each) about the Opponent's choice. Note that the model does not assume that these motives drive the participant's own choices, but rather that the participant might consider these motives as drivers of the other Player's decisions. To ensure that no other hypothetical patterns of choices across games were inferred by our participants (motives or otherwise), we compared these four canonical motives with 5,000 randomly generated pseudo-motives, which provides a conservative test of model specificity (Supplementary Results 1).

Model 2 (no structure) learns over 17 features: one for each unique (S,T) combination and one 'intercept' feature that spans the entire game space. It can thus generalize information across all games equally and treat each game as unique, but it cannot selectively generalize within a subset of games. It reflects the hypothesis that participants generalize learning about a Player only in the coarsest way possible—across all games equally—without realizing how several different games might elicit the same, or different, decisions from a Player.

Model 3 (game-type-based structure) learns over five features: one for each game type (that is, each quadrant in the game space) and one intercept. This model reflects the hypothesis that participants recognize that behaviour is consistent within a game type (for example, Stag Hunt) and thus benefit from grouping learning experiences by game type. Since Player behaviour is fully consistent within each game type (except for the Envious strategy in study 3), this strategy could be very successful in the current task. However, this model cannot generalize between game types.

In each of our three candidate models, predictions for the Player's choice are made on the basis of the weighted average predictions of each included model feature (that is, motives, games and game types, respectively):

$$V_{\text{co-operate}} = \sum_{i=1}^{k} w_i f_i\,(S,T) \tag{1}$$

where $f_i(S,T)$ is the feature's prediction at the current S and T values, expressed as 1 for co-operate and −1 for defect; w_i is the associated feature weight; i is the feature number; k is the total number of features, and $V_{\text{co-operate}}$ is the value of predicting 'co-operate'. We then apply a softmax decision rule to produce the model prediction:

$$P_{\text{co-operate}} = \frac{e^{\beta V_{\text{co-operate}}}}{1 + e^{\beta V_{\text{co-operate}}}} \tag{2}$$

where β is the inverse temperature parameter of the softmax function and $P_{\text{co-operate}}$ represents the likelihood that the Player will co-operate given the current state of the model. During learning, the Player's actual choice is compared with the model's prediction to compute a prediction error, PE:

$$\text{PE} = \text{Choice} - P_{\text{co-operate}} \tag{3}$$

Where Choice is 1 if the Player co-operated and 0 if the Player defected. The prediction error is therefore ≥0 if the Player co-operated and ≤0 if the Player defected. This sign difference is used in the next step, where each feature weight w_i is updated on the basis of the prediction error, the learning rate α and the direction of the prediction, for all i from 1 to k:

$$w_{i,t+1} = w_{i,t} + \alpha \times \text{PE} \times f_i(S,T) \tag{4}$$

where $f_i(S,T)$, as before, is 1 for co-operate and −1 for defect. Combining this sign with the sign of the prediction error allows the model to increase a feature weight if the prediction of the associated feature was correct (that is, $f_i = 1$ and PE > 0 or $f_i = -1$ and PE < 0), and conversely to decrease the weight if the feature was incorrect (when f_i and PE have opposite signs).

Model fitting. The feature weights are reset to their a priori values at the start of each task block (that is, predicting each new Player). These prior weights are free parameters that are determined when fitting the model to each participant's data, along with the learning rate α and the softmax inverse temperature β. Model 1 thus has 6 free parameters (4 feature weights, α and β), Model 2 has 19 and Model 3 has 7.

To fit the free parameters, we defined the objective function as the sum of squared errors between the model prediction and the participant's prediction weighted by confidence. Confidence-weighted predictions x are computed as follows:

$$x = \begin{cases} \frac{1+\text{confidence}}{2}, \text{co-operate} \\ \frac{1-\text{confidence}}{2}, \text{defect} \end{cases} \tag{5}$$

For example, if the participant rated 70% confidence that the Player would defect, the confidence-weighted prediction is (1 − 0.7)/2 = 0.15. To compute the sum of squared error of the model, we compared the confidence-weighted prediction with the model prediction:

$$\text{SSE} = \sum_{t=1}^{n} \left(x - P_{\text{co-operate}}\right)^2 \tag{6}$$

where t indicates trial, n indicates the total number of trials (64 in study 1, 128 in study 2) and SSE indicates the sum of squared model error. We used the Matlab function fmincon to find the parameter combination that minimizes SSE. Finally, we computed the BIC to penalize model fit by model complexity (that is, the number of free parameters), using the assumption that model error was normally distributed:

$$\text{LL} = \sum_{t=1}^{n} \log \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\left(x - P_{\text{co-operate}}\right)^2}{2\sigma^2}} \tag{7}$$

$$\text{BIC} = -2\text{LL} + (k+2) \times \log n \tag{8}$$

where σ is the standard deviation of the model errors $x - P_{\text{co-operate}}$, k is the number of feature weights in the model ($k + 2$ thus represents the number of free parameters), n is the number of observations (trials) and LL refers to log-likelihood. The BIC is an appropriate choice of model selection metric, as it allowed for reliable model selection in synthetic data (Supplementary Results 1).

Study 2 procedure. Fifty participants (12 males and 38 females; mean age, 22.2 ± 7.3 yr) played eight blocks of the Social Prediction Game in the laboratory while undergoing concurrent eye-tracking. The participants were paid for participation. The sample size was determined by finding the smallest subsets of data from study 1 in which all important effects could be reliably observed. All participants gave written informed consent before the start of the experiment. Each participant played two blocks with each Player strategy (Greedy, Risk-Averse and so on) in pseudorandom order such that each Player type occurred once in the first four blocks and once in the second four blocks. Each Player was labelled with a unique set of initials, regardless of strategy. The participants took a break between blocks 4 and 5. The task was generated using the Psychophysics Toolbox for Matlab. Otherwise, the task and computational model used in study 2 were identical to those in study 1.

Eye-tracking. Eye-tracking data were collected using a SensoMotoric Instruments iView X RED. The participants were seated comfortably and upright in a straight-backed chair 50–70 cm from the screen and were instructed to move as little as possible. The Social Prediction Game was preceded by a proprietary SensoMotoric Instruments eye-tracker calibration and validation procedure in Matlab, which was repeated if necessary until the deviation in visual angle between the validation target and the recorded gaze fixation was below two degrees. The average deviation after calibration (x, y) was (0.65°, 0.86°) for the left eye and (0.74°, 0.88°) for the right eye. The calibration and validation procedures were repeated after the break before task block 5.

The gaze data were preprocessed by converting the raw data to fixations by grouping data points on the basis of spatial and temporal distance using SensoMotoric Instruments software with the default settings. Next, all fixations outside of the bounding box of the rectangular payoff matrix on the screen were excluded. Circular regions of interest (ROIs) were then defined with 100-pixel radii around the centre of each of the eight numbers in the payoff matrix. We then computed the relative gaze to each ROI by dividing the summed fixation duration in that ROI by the total fixation duration in the entire payoff matrix. We computed relative gaze to each ROI per eye and then averaged over the two eyes to obtain the overall relative gaze to each ROI. For the $S - T$ difference scores, we subtracted relative gaze to T (summed over both T ROIs on the screen) from relative gaze to S (summed over both S ROIs).

Study 3 procedure. A total of 153 participants gave written informed consent before starting the experiment and participated on MTurk in exchange for monetary compensation. The sample size was intended to equal that of study 1. Since 5 participants did not complete the debriefing survey, only 148 participants were included in the analysis (90 males, 58 females; mean age, 36.9 ± 9.7 yr). The participants played the same Social Prediction Game as in study 1, except that there were now four Player types: Co-operative (always co-operates), Greedy (as in study 1), Risk-Averse (as in study 1) and Envious (only co-operates if own payoff ≥ other's payoff—that is, only in (S,T) pairs (7,5), (10,5) and (10,8)).

Study 4 procedure. A total of 153 participants (92 males, 61 females; mean age, 37.6 ± 10.4 yr) gave written informed consent before starting the experiment and participated on MTurk in exchange for monetary compensation. The sample size was intended to equal that of study 1. The participants first played one block of the Social Prediction Game and then played a block of the Inspection Game, with the same Player. The participants then played another block of the Social Prediction Game and the Inspection Game with a different Player. For both Players, they could thus use what they had learned about that specific Player in the Social Prediction Game to improve their decision-making in the Inspection Game. Critically, the Inspection Game is structurally distinct from all games included in studies 1–3 in that it is not symmetric (for example, mutual co-operation yields unequal payoffs). Moreover, whereas our participants only observed other Players in the Social Prediction Game, they actively participated in the Inspection Game, and their monetary reward was yoked to this game's outcomes. The Inspection Game is thus a strong test for selective generalization of learned information about the Player to a novel task.

The two Players were programmed as either Greedy or Risk-Averse in random order. These two strategies were identical to those in studies 1 and 2 for the Social Prediction Game. For the Inspection Game, the Greedy Player always shirks, and the Risk-Averse Player always works. Since there was no feedback during the Inspection Game, decision-making was driven solely by what the participant had learned during the Social Prediction Game. On each round of the Inspection Game, the participant could earn money from a base pay of US\$30 and the revenue generated by a working Employee (but not a shirking one), diminished by the wage paid to the Employee and the cost of an inspection (only if chosen). The Inspection Game therefore provides an incentive for the Employer to inspect in order to avoid paying a wage unnecessarily, but only if the Employee is shirking. The participant's earnings in one randomly selected round of the Inspection Game were paid out as a bonus with US\$1 in the game converted to US\$0.01 for the participant.

At the start of each Inspection Game block, the participant self-reported how likely they believed it was that the current Player (indicated by initials) would be working in the Inspection Game. The participants rated the Greedy Player at 41.3% likely to work and the Risk-Averse Player at 54.8%, demonstrating that inferences about others' motives were explicitly available. During the Inspection Game, a staircase procedure determined the participant's willingness to pay for inspection in the current round by raising the cost of inspecting if the participant had chosen to inspect in the previous trial and lowering it if the participant had chosen not to inspect. The willingness to pay (that is, the indifference point) was computed by averaging the cost of inspecting across the last 5 of the 15 trials of the staircase procedure. Additionally, at the end of each Inspection Game block, the participant self-reported how much they would be willing to pay for inspection (between US\$0 and US\$30).

Software and code availability. Low-level behavioural data analysis (for example, computing averages and running *t*-tests and one-way *F*-tests) was carried out in Python v.3.7.4 using the packages Numpy v.1.17.2 (ref. [65]), Pandas v.0.25.1 (ref. [66]) and Scipy v.1.3.1 (ref. [67]). The figures were created using Matplotlib v.3.1.1 (ref. [68]), Seaborn v.0.9.0 (ref. [69]), and WordCloud v.1.8.1 (ref. [70]) for Python. Mixed-effects regressions were carried out in R using the packages lme4 v.1.1–21 (ref. [71]) and lmerTest v.3.1–1 (ref. [72]). Computational modelling was performed in Matlab R2019b using the fmincon function from the Optimization Toolbox and custom code.

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

Data availability
The behavioural data analysed in this paper are available at https://github.com/jeroenvanbaar/NHB_motives_structure.

Code availability
The analysis code for this paper is available at https://github.com/jeroenvanbaar/NHB_motives_structure.

Received: 14 September 2020; Accepted: 1 September 2021;
Published online: 08 November 2021

References
1. Stewart, A. J. et al. Information gerrymandering and undemocratic decisions. *Nature* **573**, 117–121 (2019).
2. Hauser, O. P., Rand, D. G., Peysakhovich, A. & Nowak, M. A. Cooperating with the future. *Nature* **511**, 220–223 (2014).
3. Fraser, C., Riley, S., Anderson, R. M. & Ferguson, N. M. Factors that make an infectious disease outbreak controllable. *Proc. Natl Acad. Sci. USA* **101**, 6146–6151 (2004).
4. Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. How to grow a mind: statistics, structure, and abstraction. *Science* **331**, 1279–1285 (2011).
5. Seuntjens, T. G., Zeelenberg, M., Van De Ven, N. & Breugelmans, S. M. Dispositional greed. *J. Pers. Soc. Psychol.* **108**, 917–933 (2015).
6. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114**, 817–868 (1999).
7. Bolton, G. & Ockenfels, A. ERC: a theory of equity, reciprocity, and competition. *Am. Econ. Rev.* **90**, 166–193 (2000).
8. Weber, E. U., Blais, A.-R. & Betz, N. E. A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. *J. Behav. Decis. Mak.* **15**, 263–290 (2002).
9. Kahneman, D. & Tversky, A. Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 263–292 (1979).
10. Peysakhovich, A., Nowak, M. A. & Rand, D. G. Humans display a 'cooperative phenotype' that is domain general and temporally stable. *Nat. Commun.* **5**, 4939 (2014).
11. Van Lange, P. A. M. The pursuit of joint outcomes and equality in outcomes: an integrative model of social value orientation. *J. Pers. Soc. Psychol.* **77**, 337–349 (1999).
12. van Baar, J. M., Chang, L. J. & Sanfey, A. G. The computational and neural substrates of moral strategies in social decision-making. *Nat. Commun.* **10**, 1483 (2019).
13. Poncela-Casasnovas, J. et al. Humans display a reduced set of consistent behavioral phenotypes in dyadic games. *Sci. Adv.* **2**, e1600451 (2016).
14. van Baar, J. M., Klaassen, F. H., Ricci, F., Chang, L. J. & Sanfey, A. G. Stable distribution of reciprocity motives in a population. *Sci. Rep.* **10**, 18164 (2020).
15. Jara-Ettinger, J. Theory of mind as inverse reinforcement learning. *Curr. Opin. Behav. Sci.* **29**, 105–110 (2019).
16. Bridgers, S., Jara-Ettinger, J. & Gweon, H. Young children consider the expected utility of others' learning to decide what to teach. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-019-0748-6 (2019).
17. Liu, S., Ullman, T. D., Tenenbaum, J. B. & Spelke, E. S. Ten-month-old infants infer the value of goals from the costs of actions. *Science* **358**, 1038–1041 (2017).
18. Baker, C. L., Saxe, R. & Tenenbaum, J. B. Action understanding as inverse planning. *Cognition* **113**, 329–349 (2009).
19. Baker, C. L., Jara-Ettinger, J., Saxe, R. & Tenenbaum, J. B. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nat. Hum. Behav.* **1**, 0064 (2017).
20. Nihonsugi, T., Ihara, A. & Haruno, M. Selective increase of intention-based economic decisions by noninvasive brain stimulation to the dorsolateral prefrontal cortex. *J. Neurosci.* **35**, 3412–3419 (2015).
21. Wedekind, C. & Milinski, M. Cooperation through image scoring in humans. *Science* **288**, 850–852 (2000).
22. Harbaugh, W. T., Mayr, U. & Burghart, D. R. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* **316**, 1622–1625 (2007).
23. Niv, Y. et al. Reinforcement learning in multidimensional environments relies on attention mechanisms. *J. Neurosci.* **35**, 8145–8157 (2015).
24. Radulescu, A., Niv, Y. & Ballard, I. Holistic reinforcement learning: the role of structure and attention. *Trends Cogn. Sci.* **23**, 278–292 (2019).
25. Leong, Y. C., Radulescu, A., Daniel, R., Dewoskin, V. & Niv, Y. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* **93**, 451–463 (2017).
26. Collins, A. G. E. & Frank, M. J. Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition* **152**, 160–169 (2016).
27. Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B. & Botvinick, M. M. Neural representations of events arise from temporal community structure. *Nat. Neurosci.* **16**, 486–492 (2013).
28. Wilson, R. C. & Niv, Y. Inferring relevance in a changing world. *Front. Hum. Neurosci.* **5**, 189 (2012).
29. Schapiro, A. C., Turk-Browne, N. B., Norman, K. A. & Botvinick, M. M. Statistical learning of temporal community structure in the hippocampus. *Hippocampus* **26**, 3–8 (2016).
30. Gershman, S. J. & Niv, Y. Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* **20**, 251–256 (2010).
31. Kemp, C. & Tenenbaum, J. B. The discovery of structural form. *Proc. Natl Acad. Sci. USA* **105**, 10687–10692 (2008).
32. Heckathorn, D. D. The dynamics and dilemmas of collective action. *Am. Sociol. Rev.* **61**, 250–277 (1996).
33. Axelrod, R. Effective choice in the prisoner' s dilemma. *J. Conflict Resolut.* **24**, 3–25 (1980).
34. Fehr, E. & Fischbacher, U. Social norms and human cooperation. *Trends Cogn. Sci.* **8**, 185–190 (2004).
35. Goeree, J. K., Holt, C. A. & Palfrey, T. R. Risk averse behavior in generalized matching pennies games. *Games Econ. Behav.* **45**, 97–113 (2003).
36. Bernoulli, D. Exposition of a new theory on the measurement of risk. *Econometrica* **22**, 23–36 (1954).
37. Huygens, C. *The Value of All Chances in Games of Fortune* (Woodward, 1714).
38. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).
39. Hampton, A. N., Bossaerts, P. & O'Doherty, J. P. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl Acad. Sci. USA* **105**, 6741–6746 (2008).
40. Nosenzo, D., Offerman, T., Sefton, M. & Van Der Veen, A. Discretionary sanctions and rewards in the repeated inspection game. *Manage. Sci.* **62**, 502–517 (2016).
41. FeldmanHall, O. & Shenhav, A. Resolving uncertainty in a social world. *Nat. Hum. Behav.* **3**, 426–435 (2019).
42. Gershman, S. J., Pouncy, H. T. & Gweon, H. Learning the structure of social influence. *Cogn. Sci.* **41**, 545–575 (2017).
43. Shin, Y. S. & Niv, Y. Biased evaluations emerge from inferring hidden causes. *Nat. Hum. Behav.* **5**, 1180–1189 (2021).
44. Lau, T., Pouncy, H. T., Gershman, S. J. & Cikara, M. Discovering social groups via latent structure learning. *J. Exp. Psychol. Gen.* **147**, 1881–1891 (2018).

45. Jara-Ettinger, J. Theory of mind as inverse reinforcement learning. *Curr. Opin. Behav. Sci.* **29**, 105–110 (2019).
46. Collette, S., Pauli, W. M., Bossaerts, P. & O'Doherty, J. Neural computations underlying inverse reinforcement learning in the human brain. *eLife* **6**, e29718 (2017).
47. Gobet, F. et al. Chunking mechanisms in human learning. *Trends Cogn. Sci.* **5**, 236–243 (2001).
48. Thalmann, M., Souza, A. S. & Oberauer, K. How does chunking help working memory? *J. Exp. Psychol. Learn. Mem. Cogn.* **45**, 37–55 (2019).
49. Johnson, E. J., Camerer, C., Sen, S. & Rymon, T. Detecting failures of backward induction: monitoring information search in sequential bargaining. *J. Econ. Theory* **104**, 16–47 (2002).
50. Nagel, R. Unraveling in guessing games: an experimental study. *Am. Econ. Rev.* **85**, 1313–1326 (1995).
51. Rapoport, A. Individual strategies in a market entry game. *Group Decis. Negot.* **4**, 117–133 (1995).
52. Kollock, P. Social dilemmas: the anatomy of cooperation. *Annu. Rev. Sociol.* **24**, 183–214 (1998).
53. Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404 (2001).
54. Kurzban, R. & Houser, D. Experiments investigating cooperative types in humans: a complement to evolutionary theory and simulations. *Proc. Natl Acad. Sci. USA* **102**, 1803–1807 (2005).
55. Moore, T. & Zirnsak, M. Neural mechanisms of selective visual attention. *Annu. Rev. Psychol.* **68**, 47–72 (2017).
56. Caprariello, P. A., Cuddy, A. J. C. & Fiske, S. T. Social structure shapes cultural stereotypes and emotions: a causal test of the stereotype content model. *Group Process. Intergroup Relat.* **12**, 147–155 (2009).
57. McCauley, C., Stitt, C. L. & Segal, M. Stereotyping: from prejudice to prediction. *Psychol. Bull.* **87**, 195–208 (1980).
58. Nickerson, R. S. Confirmation bias: a ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **2**, 175–220 (1998).
59. Pierson, E. et al. A large-scale analysis of racial disparities in police stops across the United States. *Nat. Hum. Behav.* https://doi.org/10.1016/j.athoracsur.2014.09.078 (2020).
60. Gureckis, T. M. et al. psiTurk: an open-source framework for conducting replicable behavioral experiments online. *Behav. Res. Methods* **48**, 829–842 (2016).
61. Licht, A. N. Games commissions play: 2×2 games of international securities regulation. *Yale J. Int. Law* **24**, 61–125 (1999).
62. Bramoullé, Y. Anti-coordination and social interactions. *Games Econ. Behav.* **58**, 30–49 (2007).
63. Skyrms, B. The Stag Hunt. *Proc. Addresses Am. Phil. Assoc.* **75**, 31–41 (2001).
64. Von Neumann, J. & Morgenstern, O. *Theory of Games and Economic Behavior* (Princeton Univ. Press, 1944).
65. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
66. McKinney, W. Data structures for statistical computing in Python. In *Proc. 9th Python in Science Conf.* (eds van der Walt, S. & Millman, J.) 56–61 (2010).
67. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
68. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Engineer.* **9**, 90–95 (2007).
69. Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
70. Mueller, A. WordCloud for Python documentation, https://amueller.github.io/word_cloud/ (2020).
71. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
72. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).

Acknowledgements

We thank A. Sánchez for sharing experimental data from ref. [13]. This work was funded by NIH Centers of Biomedical Research Excellence grant no. P20GM103645 (to O.F.H). The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

J.M.v.B. and O.F.H designed the research. J.M.v.B. and W.D. performed the research. J.M.v.B., M.R.N. and O.F.H analysed the data. J.M.v.B. and O.F.H wrote the paper. J.M.v.B., M.R.N. and O.F.H edited the manuscript.

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/s41562-021-01207-4.

Correspondence and requests for materials should be addressed to Oriel FeldmanHall.

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

nature portfolio

Corresponding author(s): Oriel FeldmanHall

Last updated by author(s): Jun 30, 2021

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	Studies 1, 3, and 4: task was written in Javascript and made accessible using Psiturk 2.3.0. Study 2: task written using Psychophysics Toolbox for Matlab and eye-tracking performed using SensoMotoric Instruments iView toolbox for Matlab.
Data analysis	Basic behavioral data analysis (e.g. computing averages, running t-tests and one-way F-tests) was carried out in Python 3.7.4 using the packages Numpy 1.17.2, Pandas 0.25.1, and Scipy 1.3.1. Figures were created using Matplotlib 3.1.1 and Seaborn 0.9.0 for Python. Mixed-effects regressions were carried out in R using the packages lme4 1.1-21 and lmerTest 3.1-1. Computational modeling was performed in Matlab R2019b using the fmincon function from the Optimization Toolbox and custom code.

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

All data needed to reproduce the results in this paper will be made available on a public repository upon publication.

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	Quantitative experimental
Research sample	Study 1: American adults on Amazon Mechanical Turk Study 2: undergraduate students Study 3: American adults on Amazon Mechanical Turk Study 4: American adults on Amazon Mechanical Turk
Sampling strategy	Study 1: sample size (n = 150) was selected such that key effects from smaller pilot studies could be observed with very high statistical power. Study 2: sample size was determined by randomly sampling smaller subsets of data from study 1 such that all important effects could still be reliably observed. This was the case for n = 50. Study 3: sample size was intended to equal that of study 1 so all key effects from that study could be reliably replicated. Study 4: sample size was intended to equal that of study 1 so all key effects from that study could be reliably replicated and high power would be obtained for new, untested effects (i.e. the generalization of learning across tasks).
Data collection	Study 1: computer (online study). No-one was present besides the participant. Study 2: computer (in lab) and eye-tracker. Besides the participant, one researcher was present in the room, behind a divider. Researcher was blind to the order of the experimental conditions (within-subjects). Study 3: computer (online). No-one was present besides the participant. Study 4: computer (online). No-one was present besides the participant.
Timing	Study 1: 8 - 13 March 2019 Study 2: 15 July - 2 August 2019 Study 3: January 21 - March 24 2021 Study 4: 6 - 13 April 2020
Data exclusions	No data were excluded from analysis.
Non-participation	No participants dropped out / declined participation.
Randomization	Participants were not allocated into experimental groups (all conditions within-subjects).

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
☐	☒	Human research participants
☒	☐	Clinical data
☒	☐	Dual use research of concern

Methods

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	Study 1: 95 males, 52 females, 3 no gender indicated; mean age 35.4 ± 10.0y

Population characteristics	Study 2: 38 females, 12 males; mean age 22.2 ± 7.3y Study 3: 90 males, 58 females; mean age 36.9 ± 9.7y Study 4: 92 males, 61 females; mean age 37.6±10.4y See above for more information.
Recruitment	Study 1: Ad made available on Amazon Mechanical Turk. No self-selection bias anticipated. Study 2: Ad on the university-wide SONA system. No self-selection bias anticipated. Study 3: Ad on Amazon Mechanical Turk. No self-selection bias anticipated. Study 4: Ad on Amazon Mechanical Turk. No self-selection bias anticipated.
Ethics oversight	Brown University Institutional Review Board

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