The Paradox of Unpredictability: Trying to Be Unpredictable Makes Sequences More Predictable

Kagari Yamada*[1,2,3], Kazushi Tsutsui[1,4,5], and Kazutoshi Kudo[1]

1. Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan
2. Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology (NICT), Suita, Japan
3. Japan Society for the Promotion of Science, Tokyo, Japan
4. Graduate School of Informatics, Nagoya University, Japan
5. Center for Advanced Intelligence Project, RIKEN, Japan

*Correspondence: kagari-marines@g.ecc.u-tokyo.ac.jp

Author contributions

K.Y and K.T. designed research; K.Y. performed research; K.Y. analyzed data; and K.Y., K.T., and K.K. wrote the paper.

Declaration of interests

The authors declare no competing interest.

Acknowledgments

This work was supported by KAKENHI (22K17673, 2381231, 24H01434, 25H01237, and 25KJ1145) from the Japan Society for the Promotion of Science (JSPS) and Strategic Basic Research Programs ACT-X (JPMJAX24LA) from the Japan Science and Technology Agency. K.Y. was supported by JSPS as a JSPS Research Fellow.

Abstract

Human intuition about randomness is systematically biased. When asked to generate random sequences, people systematically generate too many alternations, yet game theory suggests that being unpredictable to an opponent requires true randomness. In our first experiment (50 Japanese adults), we show that when the goal shifts from being random to being unpredictable, participants do the opposite, generating sequences with a strong repetition bias. In a second experiment, we asked a new sample of 50 Japanese adults to predict two model-derived sequences that were empirically calibrated to the group-level statistics from the first Experiment (Competitive-like vs Random-like). We found a paradox: the repetitive sequences intended to be unpredictable were in fact more predictable than the alternating sequences intended to be random. This paradox appears to stem from an asymmetric reinforcement learning mechanism, whereby success selectively strengthens beliefs in repetition but not alternation. Our findings bridge classic cognitive bias and strategic behavior, suggesting how a fundamental learning mechanism can shape paradoxical predictability.

Introduction

Beliefs about probability play a critical role in decision-making under uncertainty. Human intuition about probability, however, is known to systematically deviate from the laws of statistics. A prominent example of this is the tendency to perceive illusory patterns in sequences of independent events. One of the most well-documented biases is the gambler's fallacy: the mistaken belief that, in a sequence of independent and identically distributed (i.i.d.) signals, observing one outcome decreases the probability of observing the same outcome again on the next trial (Benjamin, 2019). The existence of the gambler's fallacy has been confirmed in numerous studies, ranging from laboratory experiments (Barron & Leider, 2010; Benjamin et al., 2017; Rao & Hastie, 2023; Roney & Sansone, 2015) to real-world settings (Chen et al., 2016; Misirlisoy & Haggard, 2014; Suetens et al., 2016).

Closely related to how people perceive probability is the question of how they generate probabilistic sequences. When asked to generate a "random" sequence, people exhibit the same bias towards excessive alternation seen in sequence perception, systematically deviating from true randomness (Castillo et al., 2024; Guseva et al., 2023; Farmer et al., 2017; Schulz et al., 2012; Warren et al., 2018). This suggests that the gambler's fallacy reflects a deeply ingrained model of what randomness should look like.

However, when the goal of sequence generation shifts from simply mimicking randomness to the strategic aim of being unpredictable to an opponent, the problem transitions from one of cognitive bias to one of strategic interaction, as formalized in game theory. The normative solution in such competitive situations is a mixed-strategy Nash equilibrium (Nash, 1950; Nash, 1951; Reny, 1999). A mixed strategy involves randomizing one's own actions according to a specific probability distribution, thereby making one's behavior unpredictable and rendering the opponent indifferent to their own choice (Barraclough et al., 2004; Lee & Seo, 2011). In a simple binary choice task with equal payoffs, the Nash equilibrium strategy is to choose each of the two options with a probability of 50%—that is, to generate a truly random sequence. Theoretically, therefore, the goal of being unpredictable should be synonymous with the goal of being random. Building on a substantial body of work examining how closely human behavior conforms to normative solutions (Beard & Beil, 1994; Holt & Roth, 2004; Goeree & Holt, 1999; Kreps, 2018; McKelvey & Palfrey, 1992; Ochs, 1995; Roth & Erev, 1995) and how individuals reason strategically about others' behavior (Bosch-Domenech et al., 2002; Camerer et al., 2004; Ho et al., 1998; Nagel, 1995; Stahl & Wilson, 1994; Stahl & Wilson, 1995), an important open question is how fundamental cognitive biases about transition probabilities interact with strategic objectives in competitive settings.

Here, we investigate how humans generate sequences in a competitive environment where the goal is to be unpredictable to another person, and contrast this with sequences generated with the goal of being random. We show that when the goal is simply to be random, participants generate sequences with frequent alternations, replicating prior work (Castillo et al., 2024; Guseva et al., 2023; Farmer et al., 2017; Schulz et al., 2012; Warren et al., 2018). However, when the goal is to be unpredictable to an opponent, participants generate sequences with a strong bias towards repetition. Paradoxically, we then show that this competitive strategy, intended to be unpredictable, is in fact more predictable to other human players. Conversely, the sequences generated to be random, despite their statistical non-randomness, are harder for others to predict. We argue that this paradox can be traced to a fundamental mechanism of human learning: an asymmetric reinforcement process that selectively strengthens beliefs in repetitive patterns after success, but not alternating ones. This mechanism may help explain why predictors can readily build calibrated models of repetitive sequences while persistently failing to learn alternating statistics.

Results

Competitive goal reverses the human bias against repetition

In Experiment 1, participants were instructed to generate a sequence of binary choices (left or right) (Fig. 1A, B). They were randomly assigned to either a Competitive group or a Random group. The Competitive group (n = 25) was instructed to generate a sequence that would be difficult for an opponent to predict, whereas the Random group (n = 25) was instructed to generate a sequence that was as random as possible. Consistent with previous findings (Castillo et al., 2024; Guseva et al., 2023; Farmer et al., 2017; Schulz et al., 2012; Warren et al., 2018), we predicted that sequences generated by the Random group would exhibit an alternation probability higher than the 0.5 expected under a truly random sequence. Conversely, we hypothesized that the Competitive group would generate sequences with a repetition probability higher than 0.5, in order to exploit the common human bias of over-expecting alternation when predicting random-like sequences.

To test our hypotheses, we analyzed the probability that a choice would be immediately repeated (i.e., first-order transitional probability) (Fig. 1C). In line with our predictions, the Random group's repetition probability was significantly lower than chance (t (24) = -3.01, p = 0.006, *Cohen's d* (d) = -0.60, 95% confidence interval (CI) [-1.05 -0.15]), indicating a bias toward alternation. Conversely, the Competitive group's probability of repeating a choice was significantly greater than the 0.5 chance level (t (24) = 3.52, p = 0.002, d = 0.70, CI [0.24 1.17]). A direct comparison confirmed that the repetition probability was substantially higher in the Competitive group than in the Random group (t (48) = 4.44, p < 0.001, d = 1.26, CI [0.63 1.88]). Furthermore, this difference in generative strategy was not attributable to a simple preference for one side, as there was no significant difference between the groups in either directional choice bias (t (48) = -0.87, p = 0.387, d = -0.25, CI [-0.82 0.32]) (Supplementary Fig. 1A) or the overall magnitude of bias (U = 360.5, p = 0.355, r = 0.13, CI [-0.15 0.4]) (Supplementary Fig. 1B).

To confirm that these behavioral differences were not driven by differences in task engagement, we analyzed participants' self-reported ratings of how well they kept their instructions in mind throughout the experiment. 2 (group: Competitive, Random) × 10 (block: 1–10) mixed-design ANOVA on these ratings revealed no significant main effect of group (F (1, 48) = 1.48, p = 0.230, η_G^2 = 0.021), no main effect of block (F (4.71, 226.28) = 0.59, p = 0.696, η_G^2 = 0.004), and no interaction (F (4.71, 226.28) = 0.79, p = 0.550, η_G^2 = 0.005) (Supplementary Fig. 2A). This indicates that both groups were equally and consistently engaged with their respective tasks, suggesting that the observed differences in sequence generation strategies are attributable to the

experimental instructions themselves, rather than to variations in task engagement.

To further investigate how choice history influenced subsequent selections, we analyzed the conditional probability of repeating a choice given the length of the current run of identical choices (Fig. 1D). Because more than half of the participants in the Random group never generated a run of five or more identical choices, we restricted our statistical analysis to run lengths of one to four. While the two groups did not differ in their repetition probability following an alternation (i.e., a run of one, t (48) = 1.07, p = 0.288, d = 0.30, CI [-0.27 0.88]), a significant gap emerged as runs grew longer. The Competitive group was significantly more likely to continue a run than the Random group after two (t (48) = 5.40, $p < 0.001$, d = 1.53, CI [0.88 2.18]), three (t (48) = 5.73, $p < 0.001$, d = 1.62, CI [0.96 2.28]), and four (t (43) = 3.48, p = 0.002, d = 1.04, CI [0.40 1.69]) consecutive identical choices. When compared to the 0.5 chance level, the Competitive group showed a significantly higher probability of repetition after a run of two (t (24) = 4.23, p = 0.001, d = 0.85, CI [0.36 1.33]), consistent with exploiting streak expectations. In contrast, the Random group exhibited a bias against repetition, as their probability of repeating a choice was significantly lower than chance after runs of two (t (24) = -3.68, p = 0.002, d = -0.74, CI [-1.20 -0.27]), three (t (24) = -5.96, $p < 0.001$, d = -1.19, CI [-1.74 -0.65]), and four (t (19) = -3.34, p = 0.005, d = -0.75, CI [-1.28 -0.21]).

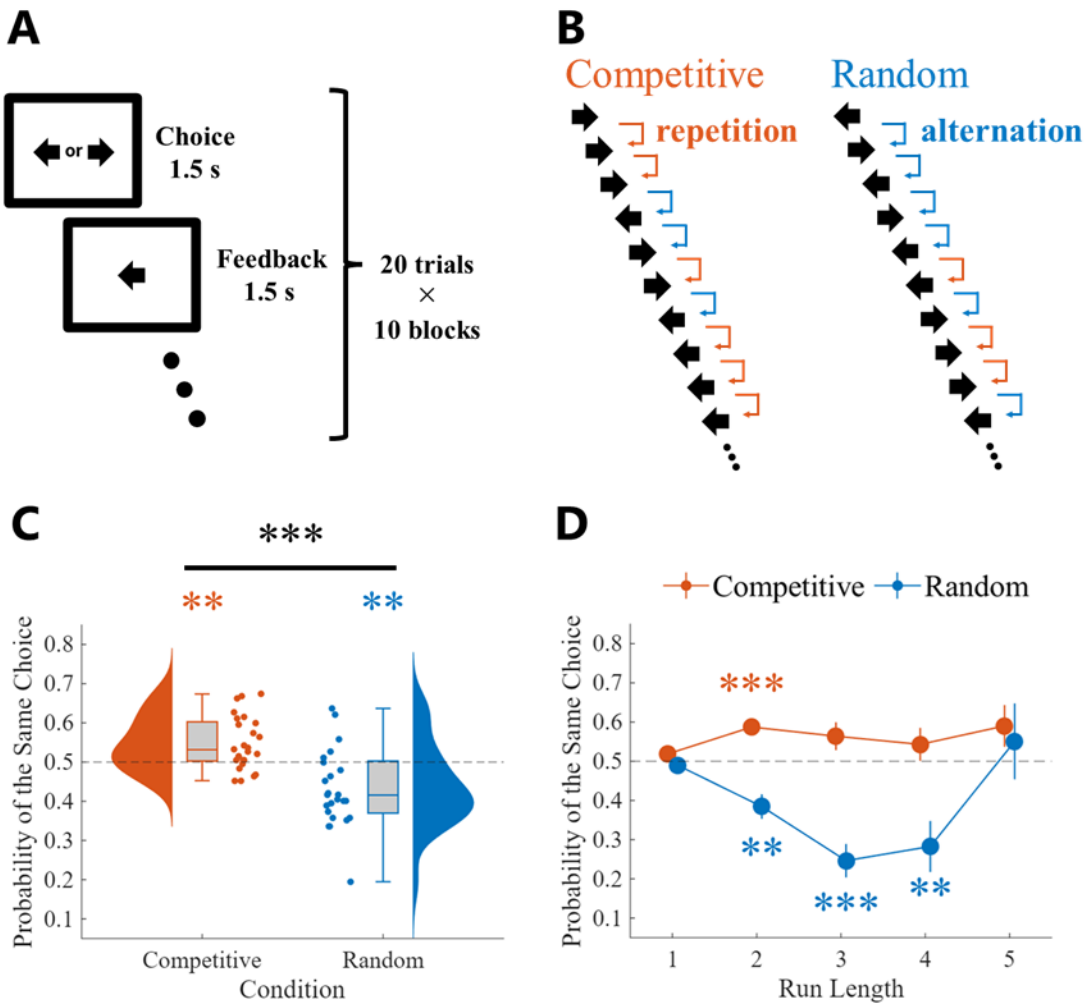

Figure 1. Competitive goal reverses the human bias against repetition.
(A) Example of procedures in Experiment 1. Each of the 10 blocks consisted of 20 binary choices. Following each block, participants' awareness of the "unpredictable" in the Competitive group or "random" in the Random group task instruction was checked. (B) Example segments of generated sequences and definition of choice transitions. The sequences illustrate the tendency for repetition in the Competitive condition and alternation in the Random condition. A repetition (red) occurs when a choice is the same as the one preceding it (for example, left following left). An alternation (blue) occurs when a choice is different from the one preceding it (for example, right following left). (C) Participants in the Competitive group repeated choices more often than those in the Random group. The distribution of choice repetition probabilities is shown for the Competitive (red) and Random (blue) groups. The shaded area of each half-violin plot represents the probability density of the data. The embedded box plot indicates the median (central line) and the interquartile range, while the black circles represent

individual participants. The horizontal dashed line indicates the chance level (0.5). A black asterisk denotes a significant difference between the groups, while red and blue asterisks denote a significant difference for each respective group from the chance level. (D) The plot shows the mean probability of repeating a choice given the number of immediately preceding identical choices (run length). Data are presented for both the Competitive (red line) and Random (blue line) groups. Error bars represent ±1 standard error of the mean (SEM). The dashed horizontal line indicates the chance level of 0.5. Red and blue asterisks denote a significant difference for each respective group from the chance level.

Significant effects are indicated by an asterisk. $^{**}p < 0.01$, $^{***}p < 0.001$.

Striving for unpredictability paradoxically increases human-to-human predictability

To formally test whether the generative bias observed in Experiment 1 produces sequences that are indeed difficult for humans to predict, we conducted Experiment 2 that required participants to forecast these sequences. Rather than using the raw data from a single participant from Experiment 1, which might be idiosyncratic (Jokar & Mikaili, 2012; Schulz et al., 2012), we generated two empirically calibrated synthetic sequences designed to capture the quintessential statistical properties of each group. To do this, we constructed a generative model for each group (Competitive and Random). The model's parameters were the group-average conditional probabilities of choice repetition, contingent on the length of the current run, as empirically determined in Experiment 1. We then ran a large-scale simulation (1 billion iterations) and selected the one sequence for each group whose conditional probability profile most closely matched that of the original group-average data (see Methods). These two sequences served as the stimuli to be predicted in Experiment 2.

Before testing predictability, we verified that the two simulated synthetic sequences used in Experiment 2 were closely matched in information content. The marginal left/right frequencies were balanced in both sequences (Competitive: Left 101/Right 99; Random: Left 100/Right 100). Accordingly, the symbol entropy was maximal for both ($H(symbol)$ = 1.000 bit in each case). First-order structure was also essentially matched: the Markov entropy was near maximal and differed only trivially between sequences (Competitive: $H(X_t|X_{t-1})$ = 0.986 bit; Random: $H(X_t|X_{t-1})$ = 0.988 bit), and the entropy of the repetition/alternation label was similarly high (Competitive: 0.986 bit; Random: 0.988 bit). Together, these checks indicate that marginal and first-order statistics do not favor either sequence.

Going beyond first-order, conditioning on run length yielded only modest additional reductions in uncertainty for both sequences. The run-conditioned symbol entropy remained high (Competitive: $H(X|run)$ = 0.948 bit; Random: $H(X|run)$ = 0.983 bit), and the run-conditioned repetition/alternation entropy was likewise high (Competitive: $H(rep/alt|run)$ = 0.962 bit; Random: $H(rep/alt|run)$ = 0.973 bit). As a single summary index of higher-order information beyond first order, the additional reduction was only on the order of hundredths of a bit (Competitive: $\Delta H(run|prev)$ = 0.024 bit; Random: $\Delta H(run|prev)$ = 0.015 bit). Thus, while slightly lower entropies for the Competitive sequence suggest that higher-order regularities may exist, their magnitude is small relative to the overall 1-bit uncertainty.

In Experiment 2, we tasked a new set of participants with predicting the generated sequences on a trial-by-trial basis (Fig. 2A, B). Participants were randomly assigned to predict the sequence generated by either the Competitive model (n = 25) or the Random model (n = 25). None of these participants had taken part in Experiment 1. To avoid biasing their predictions (Rao & Hastie, 2023), participants were provided with no information regarding the origin or generative process of the sequences they were to predict (for instance, how or by whom they were created).

We first analyzed the overall prediction accuracy (i.e., success rate) for each of the two generated sequences (Fig. 2C). We found that the sequence generated by the Competitive model was more predictable than the sequence from the Random model (t (48) = 2.02, p = 0.049, d = 0.57, CI [-0.01 1.15]). However, when tested against the 0.5 chance level, neither group's performance was reliably different from guessing. Neither group's prediction accuracy was significantly different from the 0.5 chance level (Competitive: t (24) = 1.49, p = 0.150, d = 0.30, CI [-0.12 0.72]; Random: t (24) = -1.39, p = 0.178, d = -0.28, CI [-0.70 0.14]).

To ensure that these differences in predictability were not confounded by varying levels of task engagement, we analyzed participants' self-reported engagement ratings. 2 (group: predicting Competitive, predicting Random) × 10 (block: 1–10) mixed-design ANOVA on these ratings found no significant main effect of group (F (1, 48) = 0.40, p = 0.528, η_G^2 = 0.006), no main effect of block (F (6.56, 314.79) = 0.45, p = 0.862, η_G^2 = 0.003), and no interaction (F (6.56, 314.79) = 0.77, p = 0.605, η_G^2 = 0.005) (Supplementary Fig. 2B). This indicates that participants in both groups remained equally and consistently engaged throughout the task, suggesting that the observed differences in their predictive behavior are attributable to the statistical properties of the sequences they were forecasting.

To understand the basis for this accuracy difference, we next analyzed participants'

trial-by-trial prediction strategies (Fig. 2D). Specifically, we examined the probability that a participant's choice would match the previous correct answer, which reflects their expectation of a repetition. This revealed that participants successfully learned the statistical structure of the Competitive sequence. They were significantly more likely than chance to predict that a choice would repeat (t (24) = 3.93, p < 0.001, d = 0.79, CI [0.31 1.26]). In stark contrast, participants predicting the Random sequence failed to learn its alternating nature; their probability of predicting a repetition was indistinguishable from chance (t (24) = -0.19, p = 0.848, d = -0.04, CI [-0.45 0.37]). Consequently, the tendency to expect and predict a repetition was significantly stronger for the group predicting the Competitive sequence (t (48) = 2.71, p = 0.009, d = 0.77, CI [0.18 1.36]).

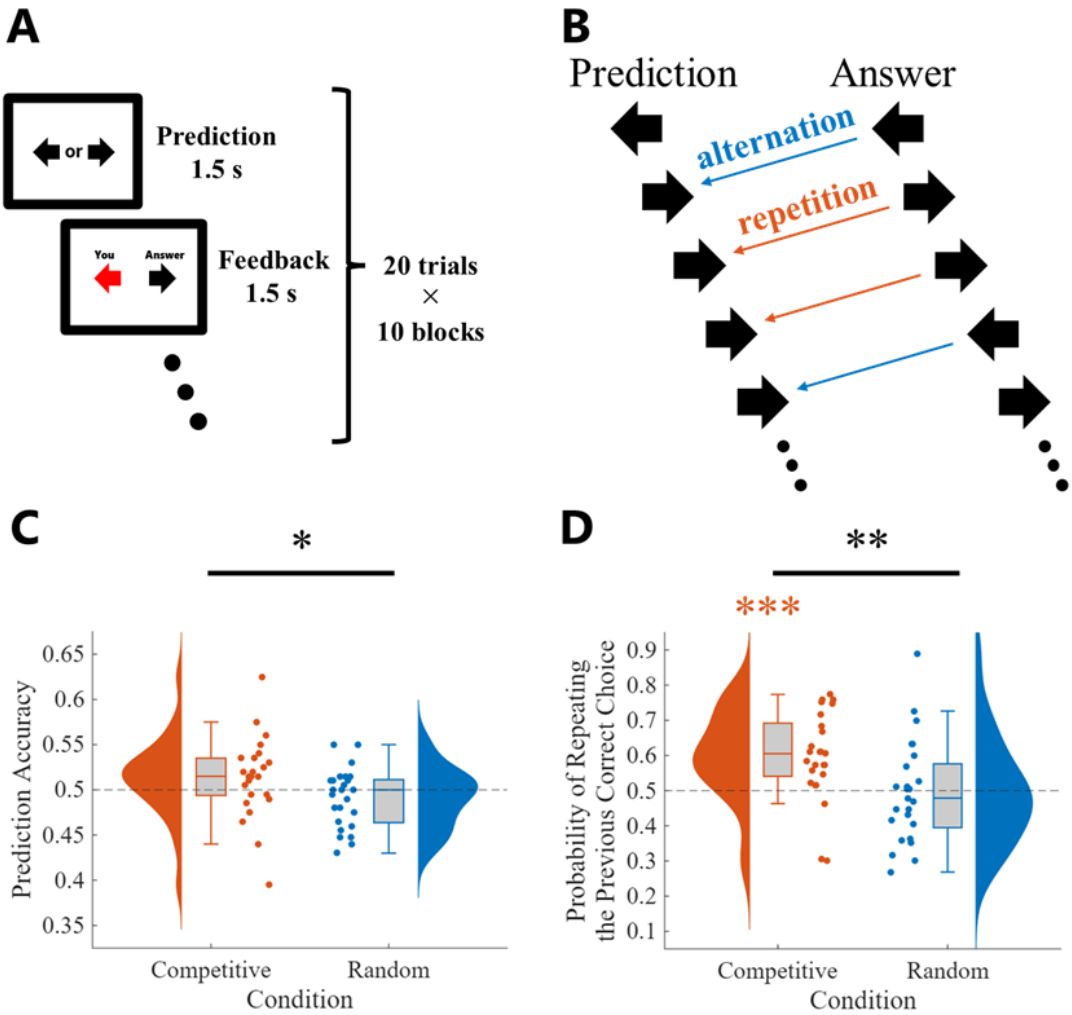

Figure 2. Striving for unpredictability paradoxically increases human-to-human predictability.

(A) Example of procedures in Experiment 2. Each of the 10 blocks consisted of 20

binary choices to predict a target sequence. Following each block, participants rated their level of engagement with the task.

(B) Definition of prediction strategies based on the preceding correct answer. A repetition strategy (red) involves predicting that the next choice will be the same as the previous correct answer. An alternation strategy (blue) involves predicting that the next choice will be different from the previous correct answer.

(C) The Competitive sequence was more predictable than the Random sequence. The distribution of prediction accuracy is shown for the Competitive and Random sequences. The shaded area of each half-violin plot represents the probability density of the data. The embedded box plot indicates the median (central line) and the interquartile range, while the black circles represent individual participants. The horizontal dashed line indicates the chance level (0.5). A black asterisk denotes a significant difference between the groups.

(D) Participants learned to expect repetitions in the Competitive sequence but not the Random sequence. The probability that participants predicted a repetition of the previous correct item in the sequence is shown for the Competitive and Random sequences. The shaded area of each half-violin plot represents the probability density of the data. The embedded box plot indicates the median (central line) and the interquartile range, while the black circles represent individual participants. The horizontal dashed line indicates the chance level (0.5). A black asterisk denotes a significant difference between the groups, while red asterisk denotes a significant difference from the chance level.

Significant effects are indicated by an asterisk. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Humans readily learn repetitive patterns but fail to learn alternating ones

We further explored these prediction strategies by examining how participants' expectation of repetition was modulated by the recent history of the sequences (Fig. 3A). Because the maximum run of identical items in the Random sequence was five, we limited this analysis to run lengths of one to five. For the Competitive sequence, participants' expectations correctly tracked its streaky nature; their probability of predicting a repetition was significantly higher than chance after observing runs of one (t (24) = 2.33, p = 0.048, d = 0.47, CI [0.03 0.90]), two (t (24) = 3.66, p = 0.003, d = 0.73, CI [0.26 1.20]), four (t (24) = 2.17, p = 0.050, d = 0.43, CI [0.00 0.87]), and five (t (24) = 3.98, p = 0.003, d = 0.80, CI [0.32 1.27]) identical items. In stark contrast, participants predicting the Random sequence showed no evidence of learning its statistical properties. Their expectation of repetition remained statistically

indistinguishable from the 0.5 chance level, regardless of the run length (one: t (24) = 0.53, p = 0.802, d = 0.11, CI [-0.31 0.52]; two: t (24) = -0.23, p = 0.820, d = -0.05, CI [-0.46 0.37]; three: t (24) = -1.61, p = 0.341, d = -0.32, CI [-0.75 0.10]; four: t (24) = -0.47, p = 0.803, d = -0.09, CI [-0.51 0.32]; five: t (24) = -1.54, p = 0.341, d = -0.31, CI [-0.73 0.11]). This fundamental difference in learned expectations led to a significant divergence in prediction strategies between the sequences, particularly after observing runs of two (t (48) = 2.63, p = 0.029, d = 0.74, CI [0.15 1.33]), three (t (48) = 2.42, p = 0.032, d = 0.69, CI [0.10 1.27]), and five (t (48) = 3.17, p = 0.013, d = 0.90, CI [0.30 1.49]) items, where those predicting the Competitive sequence had a much higher expectation of repetition.

Next, we assessed how well participants' internal models were calibrated to the true statistical properties of the sequences. To do this, we compared participants' conditional expectation of repetition against the actual conditional probability of repetition in the sequences at each run length (Fig. 3B). Participants predicting the Competitive sequence appeared to be remarkably well-calibrated; their expectation of repetition did not significantly differ from the true probabilities of the sequence at any run length (one: t (24) = 1.68, p = 0.349, d = 0.34, CI [-0.09 0.76]; two: t (24) = 1.22, p = 0.349, d = 0.24, CI [-0.18 0.66]; three: t (24) = 0.18, p = 0.859, d = 0.04, CI [-0.38 0.45]; four: t (24) = 1.11, p = 0.349, d = 0.22, CI [-0.20 0.64]; five: t (24) = 1.43, p = 0.349, d = 0.29, CI [-0.14 0.71]). In contrast, those predicting the Random sequence were systematically miscalibrated. They significantly underestimated the sequence's true anti-streaky nature, over-predicting the likelihood of a repetition following runs of two (t (24) = 2.41, p = 0.040, d = 0.48, CI [0.04 0.92]), three (t (24) = 2.76, p = 0.040, d = 0.55, CI [0.11 1.00]), and four (t (24) = 2.47, p = 0.040, d = 0.49, CI [0.06 0.93]) items. Intriguingly, this failure of learning occurred even though the Random sequence contained a more informative statistical signal, that is, its actual repetition probabilities deviated further from 0.5 than those of the Competitive sequence (Fig. 1D). This suggests that human learners can more effectively model the noisy statistics of repetition from a competitive agent than the alternating statistics of a generator trying to be random.

Finally, to examine the time course of this model-building, we analyzed how participants' expectation of repetition evolved on a block-by-block basis (Fig. 3C). This revealed a rapid learning process for the Competitive sequence. Although participants' predictions were at chance level in the first block (t (24) = -0.62, p = 0.540, d = -0.12, CI [-0.54 0.29]), their expectation of repetition rose quickly (3rd block: t (24) = 2.70, p = 0.021, d = 0.54, CI [0.10 0.98]; 4th block: t (24) = 3.03, p = 0.015, d = 0.61, CI [0.15 1.06]) and was significantly above 0.5 in most subsequent blocks (e.g., 9th block: t (24)

= 4.23, p = 0.003, d = 0.85, CI [0.36 1.33]; 10th block: t (24) = 3.53, p = 0.006, d = 0.71, CI [0.24 1.17]). Conversely, participants predicting the Random sequence showed no learning at any point; their expectation of repetition never deviated from chance level across all 10 blocks (e.g., 1st block: t (24) = -0.76, p = 0.972, d = -0.15, CI [-0.57 0.26]; 9th block: t (24) = 0.04, p = 0.972, d = 0.01, CI [-0.41 0.42]; 10th block: t (24) = 0.69, p = 0.972, d = 0.14, CI [-0.28 0.55]). As a result, the prediction strategies of the two groups, initially identical (1st block: t (48) = 0.16, p = 0.876, d = 0.04, CI [-0.52 0.61]), significantly diverged over the course of the experiment (e.g., 9th block: t (48) = 3.20, p = 0.025, d = 0.90, CI [0.31 1.50]).

To formally test this divergence, we collapsed blocks into two levels (Block 1 vs. the mean of Blocks 2–10) and ran a 2 (Sequence: Competitive vs. Random; between-subjects) × 2 (Block: early vs. late; within-subjects) mixed ANOVA on the expectation-of-repetition measure. This revealed a significant main effect of Block (F (1, 48) = 13.45, p < 0.001, η_G^2 = 0.058) and, critically, a Block × Sequence interaction (F (1, 48) = 5.81, p = 0.020, η_G^2 = 0.026), with no main effect of Sequence (F (1, 48) = 2.21, p = 0.143, η_G^2 = 0.035). Simple-effects showed that expectation of repetition increased from Block 1 to later blocks in the Competitive sequence (F (1, 24) = 17.93, p < 0.001, η_G^2 = 0.163), but not in the Random sequence (F (1, 24) = 0.82, p = 0.375, η_G^2 = 0.006).

Taken together, these results suggest that human predictors can rapidly learn and build a calibrated internal model of sequences with repetitive dependencies, but tend to persistently fail to learn—and instead systematically misperceive—sequences with alternating statistics.

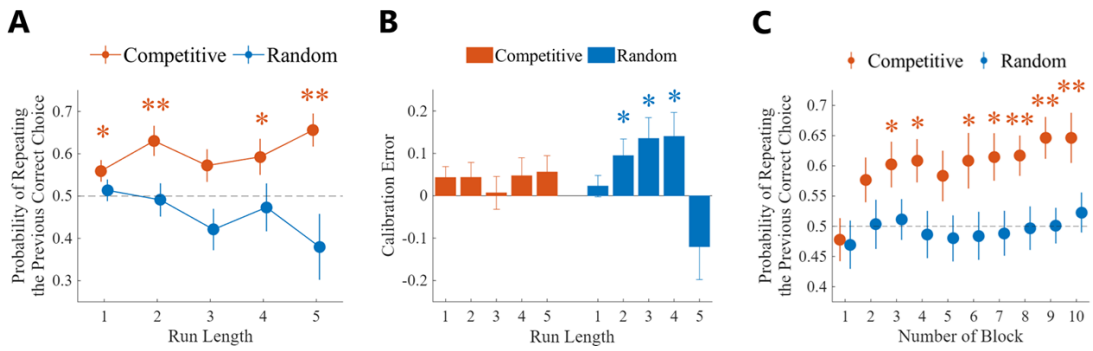

Figure 3. Humans readily learn repetitive patterns but fail to learn alternating ones. (A) Divergent learning of sequence structure. Participants' expectation of repetition is plotted as a function of the sequences run length for the groups predicting the Competitive (red line) and Random (blue line) sequences. Error bars represent ±1 SEM. The dashed horizontal line indicates the chance level of 0.5. Asterisks denote a significant difference from the chance level.

(B) Calibration of internal models to sequence statistics. The plot shows the mean calibration error (the difference between participants' subjective expectation of repetition and the actual probability of repetition), conditioned on run length. A value of zero indicates perfect calibration. Data are shown for the groups predicting the Competitive (red bars) and Random (blue bars) sequences. Error bars represent ±1 SEM. Asterisks denote that the calibration error is significantly different from zero (perfect calibration).

(C) Rapid learning of statistics of repetition but a persistent failure to model alternating patterns. The evolution of participants' expectation of repetition is shown across the 10 experimental blocks. The mean probability of predicting a repetition is plotted for the groups predicting the Competitive (red) and Random (blue) sequences. Points represent group means for each block; error bars are ±1 SEM. The dashed line at 0.5 indicates the chance level. Asterisks denote a significant difference from the chance level. Significant effects are indicated by an asterisk. $*p < 0.05$, $**p < 0.01$.

Asymmetric reinforcement learning shapes prediction strategies

To elucidate the learning mechanisms driving participants' predictions, we examined how they updated their strategy based on trial-by-trial feedback. We defined two simple strategies—predicting a repetition or predicting an alternation—and analyzed the probability of staying with the same strategy on the next trial as a function of the previous outcome (correct or incorrect).

A 2 (Sequence Type: Competitive vs. Random) × 2 (Previous Strategy: Repetition vs. Alternation) × 2 (Previous Outcome: Correct vs. Incorrect) mixed-design ANOVA revealed a complex learning process, indicated by several significant interactions. The most crucial of these was a significant interaction between the previously used strategy and its outcome (F (1, 48) = 5.19, p = 0.027, η_G^2 = 0.031), which suggested a fundamental asymmetry in how participants learned from feedback (Fig. 4A). Post-hoc tests showed that participants applied a clear "win-stay, lose-shift" rule, but only to the repetition strategy. Specifically, after a correct outcome, the probability of staying with the repetition strategy was significantly higher than after an incorrect outcome (F (1, 48) = 6.84, p = 0.012, η_G^2 = 0.048). In stark contrast, this reinforcement did not apply to the alternation strategy; correctly predicting an alternation did not significantly increase their likelihood of predicting an alternation again (F (1, 48) = 1.96, p = 0.168, η_G^2 = 0.020). The three-way interaction was not significant (F (1, 48) = 0.07, p = 0.794, η_G^2 = 0.000), indicating that this asymmetric learning rule operated irrespective of the sequence being predicted.

However, the overall context of the sequence still mattered. We found an interaction between Sequence Type and Previous Strategy (F (1, 48) = 6.60, p = 0.013, η_G^2 = 0.066) (Fig. 4B). Post-hoc tests showed that when predicting the Competitive sequence, participants were significantly more likely to stick with a repetition strategy than an alternation strategy (F (1, 24) = 13.2, p = 0.001, η_G^2 = 0.183), reflecting their successful learning of the sequence's structure. No such strategic preference was found for the Random sequence (F (1, 24) = 0.18, p = 0.672, η_G^2 = 0.005). Furthermore, a significant interaction between Sequence Type and Previous Outcome (F (1, 48) = 6.69, p = 0.013, η_G^2 = 0.020) confirmed that the "win-stay" pattern was also specific to the Competitive sequence (Fig. 4C). Post-hoc tests revealed that after a correct outcome, participants predicting the Competitive sequence were significantly more likely to stay with their chosen strategy (F (1, 24) = 4.50, p = 0.045, η_G^2 = 0.028), whereas this was not the case for those predicting the Random sequence (F (1, 24) = 2.36, p = 0.138, η_G^2 = 0.013).

These results offer a potential mechanistic explanation for our main findings, suggesting that participants' learning was driven by a simple, asymmetric reinforcement process that selectively strengthened the belief in repetition. This process may have made the repetitive Competitive sequence learnable while leaving the alternating Random sequence difficult to learn.

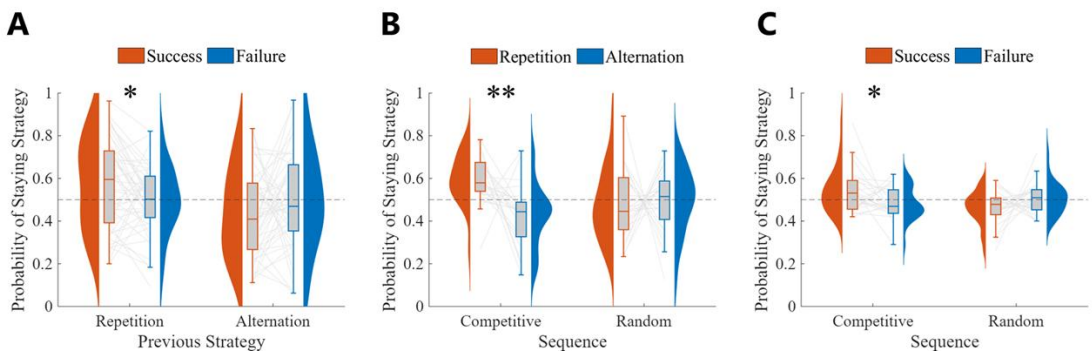

Figure 4. Asymmetric reinforcement learning shapes prediction strategies.
(A) Reinforcement is selective for repetition strategies. The plot shows the probability of staying with a given strategy (Repetition or Alternation) as a function of the previous trial's outcome (Success or Failure), collapsed across both sequence types. The significant difference following a successful repetition, but not a successful alternation, indicates that a "win-stay" rule was applied asymmetrically.
(B) A strategic preference for repetition emerges for the Competitive sequence. The plot shows the probability of staying with a repetition versus an alternation strategy for each sequence type, collapsed across outcomes. The preference to stick with a repetition strategy is evident only for the Competitive sequence.

(C) The “win-stay” learning rule is applied only when predicting the competitive sequence. The plot shows the overall probability of staying with a strategy as a function of the previous outcome for each sequence type, collapsed across strategies. The tendency to stay with a strategy after a success is present only for the Competitive sequence.

In all panels, the shaded areas represent the probability density of the data (violin plot), and embedded box plots show the median and interquartile range. Thin grey lines connect data from the same participant between conditions.

Significant effects are indicated by an asterisk. *$p < 0.05$, **$p < 0.01$.

Model-based analysis reveals asymmetric learning rates favoring repetition

To test whether the asymmetric reinforcement seen at the strategy level reflects a latent learning mechanism, we fit a simple reinforcement-learning model to trial-by-trial predictions, allowing the learning rate to differ depending on whether participants had just repeated or alternated their prediction (see Methods).

Across participants, the learning rate following a repetition choice was reliably higher than the learning rate following an alternation choice ($Z = 2.99$, $p = 0.001$, $r = 0.42$, CI [0.16 0.63]) (Fig. 5A). The same pattern held within each sequence context: when predicting the Competitive sequence ($Z = 1.86$, $p = 0.032$, $r = 0.37$, CI [-0.03 0.67]), and when predicting the Random sequence ($Z = 2.05$, $p = 0.020$, $r = 0.41$, CI [0.02 0.69]) (Fig. 5B). By contrast, the absolute magnitude of the learning rates did not differ between sequences (for the repetition learning rate: $U = 390$, $p = 0.135$, $r = 0.21$, CI [-0.07 0.46]; for the alternation learning rate: $U = 373$, $p = 0.244$, $r = 0.17$, CI [-0.12 0.42]) (Fig. 5C). Together, these findings indicate that feedback selectively strengthens a tendency to repeat, and that this asymmetry is sequence-general, rather than tied to a particular stimulus stream.

We also examined two auxiliary components of the choice rule. First, the baseline response bias (intercept) did not differ from zero in either sequence (Competitive: $Z = -0.26$, $p = 0.798$, $r = -0.05$, CI [-0.44 0.35]; Random: $Z = -0.87$, $p = 0.382$, $r = -0.17$, CI [-0.53 0.24]), and it did not differ between sequences ($U = 338$, $p = 0.628$, $r = 0.07$, CI [-0.21 0.34]) (Supplementary Fig. S3A). Second, the tendency to stick with the previous strategy (recency/inertia) was likewise not different from zero (Competitive: $Z = -0.34$, $p = 0.737$, $r = -0.07$, CI [-0.45 0.34]; Random: $Z = -1.74$, $p = 0.083$, $r = -0.35$, CI [-0.65 0.06]) and did not differ between sequences ($U = 366$, $p = 0.304$, $r = 0.15$, CI [-0.14 0.41]) (Supplementary Fig. S3B). Thus, the model-based evidence points specifically to an asymmetric learning-rate profile—higher learning rates for repetitions than for

alternations—as the mechanism shaping human prediction strategies, dovetailing with the behavioral selectivity.

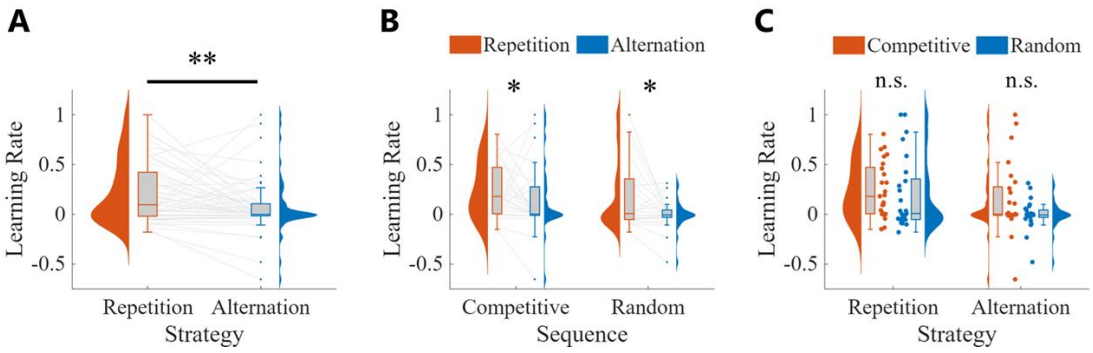

Figure 5. Model-based fits reveal a repetition-favoring learning rates.
(A) Participant-wise distributions of the estimated learning rates following repetition versus alternation choices, collapsed across sequences. The learning rate after choosing to repeat was higher than the learning rate after choosing to alternate. The shaded areas represent the probability density of the data (violin plot), and embedded box plots show the median and interquartile range. Thin grey lines connect data from the same participant between strategies.
(B) The same learning-rate asymmetry held within each sequence context (Competitive and Random). The shaded areas represent the probability density of the data (violin plot), and embedded box plots show the median and interquartile range. Thin grey lines connect data from the same participant between strategies.
(C) The absolute magnitude of each learning rate did not differ between sequences. The embedded box plot indicates the median (central line) and the interquartile range, while the black circles represent individual participants. The horizontal dashed line indicates the chance level.
Significant effects are indicated by an asterisk. *$p < 0.05$, **$p < 0.01$.

Discussion

This study aimed to elucidate the cognitive biases and strategic considerations that shape human sequence generation and prediction. We found a stark dissociation: when explicitly directed to be random, participants exhibited an alternation bias, yet when directed to be unpredictable to an opponent—a goal that is theoretically synonymous with randomness—they showed a strong repetition bias. The patterns of sequence generation observed in our study confirm that different instructional frames elicit

distinct behaviors, one rooted in a deep-seated cognitive bias and the other in a simple competitive strategy.

The excessive alternation observed when participants were instructed to generate a random sequence robustly replicates a vast body of literature (Castillo et al., 2024; Farmer et al., 2017; Guseva et al., 2023; Schulz et al., 2012; Warren et al., 2018). This clearly indicates that their default concept of randomness is not that of statistically independent events (where the probability of repetition is 0.5), but is instead distorted by a belief in "the law of small numbers" (Tversky & Kahneman, 1971). Participants, in an attempt to generate what they believe a random sequence should look like—one where balance is maintained over short intervals and repetitions are avoided—end up generating sequences that are, in fact, statistically biased (Benjamin et al., 2017; Rabin, 2002; Rabin & Vayanos, 2010; Rapoport & Budescu, 1997).

Conversely, interpreting the repetition bias observed under competitive instructions as a simple cognitive bias (for example, a "hot-hand bias" or "positive recency"; Croson & Sundali, 2005; Oskarsson et al., 2009; Suetens et al., 2016) is likely inappropriate. If participants were consciously aiming for the game-theoretic optimal strategy of true randomness (Nash, 1950; Nash, 1951; Reny, 1999), they should have exhibited the same alternation bias as the Random group. A more plausible account is that generators anticipated that their opponents would expect alternation (due to the well-known gambler's fallacy; Barron & Leider, 2010; Benjamin et al., 2017; Rao & Hastie, 2023; Roney & Sansone, 2015) and strategically chose to repeat choices to exploit this expectation. This type of strategic thinking, based on reasoning about others' thoughts, is formalized as level-k theory, which explains systematic deviations from Nash equilibrium (Bosch-Domenech et al., 2002; Camerer et al., 2004; Ho et al., 1998; Nagel, 1995; Stahl, & Wilson, 1994; Stahl, & Wilson, 1995).

One of the most intriguing findings of this study is the paradox that the sequences strategically intended to be unpredictable were, in fact, more predictable than those intended to be random. The ability to predict the next element in a sequence is a form of statistical learning (Friston, 2010; Frost et al., 2015; Tishby & Polani, 2010). The sequences generated to be unpredictable contained a clear statistical structure—a higher-than-chance probability of repetition. Our results offer insights into a potential mechanism by which participants learned this structure, suggesting they employed an asymmetric reinforcement learning rule that selectively strengthened their tendency to predict a repetition after it was successful. This "win-stay" rule for repetition strategies may explain their ability to build a calibrated model of the Competitive sequence.

In contrast, the sequences generated to be random were not learned by predictors,

despite their alternation rate deviating from 50% and even more so than the competitive sequences. This fact suggests the existence of an asymmetric baseline in human probabilistic sequence learning. The brain is thought to update its internal models based on prediction errors—the mismatch between expected and actual outcomes (Clark, 2013; Friston & Kiebel, 2009; Huang & Rao, 2011; Spratling, 2017)—and possesses powerful statistical learning mechanisms to extract regularities such as item frequencies and, crucially, transition probabilities between trials (Armstrong et al., 2017; Fang et al., 2023; Foucault & Meyniel, 2021; Maheu et al., 2019; Santolin & Saffran, 2018). Our mechanistic analysis provides direct behavioral evidence supporting the existence of this asymmetry: participants failed to reinforce the alternation strategy, even when it was successful. This finding supports the hypothesis that the brain's baseline expectation for a binary sequence is not 50% but is already biased towards alternation. Therefore, when observing a sequence that alternates, the prediction error may not be large enough to drive substantial updating of the internal model. This hypothesis is plausible, given that humans often expect repetition rates lower than chance in many contexts (Barron & Leider, 2010; Benjamin et al., 2017; Rao & Hastie, 2023; Roney & Sansone, 2015).

While this study offers new insights into human sequence generation and perception, several limitations should be noted. First, the application of level-k theory is merely an inference based on behavioral data about participants' beliefs or reasoning abilities (Jin, 2021). It is possible that some participants engaged in even deeper meta-reasoning, anticipating that an opponent would predict their exploitation of the gambler's fallacy and thus reverting to an alternating strategy (a "double bluff"). The fact that the deviation from chance was smaller for the competitive sequences than for the random sequences does not rule out this possibility. However, studies suggest that most participants do not engage in such deep strategic thinking (Camerer et al., 2004; Costa-Gomes & Crawford, 2006). Second, there are alternative explanations for the failure to learn the random sequence. The difficulty may have arisen from the non-stationary nature of its transition probabilities; the probability of repetition changed as a function of run length, from approximately 50% after one repetition to below 30% after three or four. At the same time, when we quantified the information content of the two empirically calibrated sequences, marginal and first-order statistics were closely matched, and even after conditioning on run length the additional information remained small. Thus, while a contribution of higher-order structure cannot be entirely ruled out, it is likely modest. Moreover, humans have been shown to learn higher-order transition probabilities (Remillard, 2008; Remillard, 2010; Remillard & Clark, 2001).

Research on perceptual biases in randomness, generative biases, and strategic gameplay has largely progressed in parallel. By creating a unique paradigm where sequences generated by one set of humans are predicted by another, our study directly bridges these fields. Our findings suggest how a fundamental mechanism—asymmetric reinforcement learning—could potentially explain the link between these domains. Our work suggests that the alternation bias in sequence generation, often known as the gambler's fallacy, is not merely an isolated cognitive error but a manifestation of a more fundamental, skewed internal model of 'chance' that shapes these learning processes. This insight has profound implications for understanding human perception, judgment, and strategic behavior in real-world environments involving risk.

Methods

Participants

Fifty individuals (30 female, 19 male, 1 non-binary; mean age ± SD = 30.9 ± 9.54 years, age range: 18–57 years) participated in Experiment 1. The sample size was determined a priori using G*Power (Faul et al., 2007). Based on an effect size of d = 0.76 from a pilot study for the difference between the probability of repeating a choice and the chance level, we determined that 25 participants were required for the Competitive group to achieve 95% power for a one-sample t-test against the chance level ($\alpha = 0.05$). We therefore recruited 25 participants for the Competitive group and a matched sample of 25 for the Random group.

Fifty new individuals (25 female, 25 male; mean age ± SD = 28.6 ± 9.02 years, age range: 18–57 years) participated in Experiment 2. The sample size was set to 25 participants per condition to match the sample size of Experiment 1. None of these participants had taken part in Experiment 1.

According to self-reports, all participants had normal or corrected-to-normal vision. This study was approved by the Ethical Review Committee for Experimental Research Involving Human Subjects at the Graduate School of Arts and Sciences, the University of Tokyo. All participants provided informed consent via an online form before participating in the study.

Apparatus

All experiments were conducted using the online experiment tool Gorilla Experiment Builder (https://gorilla.sc). The temporal accuracy of the tool has been validated in previous studies (Anwyl-Irvine et al., 2021; Anwyl-Irvine et al., 2020). The participants

performed the experiment in a web browser on their personal computer.

Experiment 1: Sequence Generation Task

Participants first received instructions on how to perform the task. They then completed a training session of 10 trials to familiarize themselves with the task procedure. Following the training, they were given a specific instruction for the main session. Participants in the Competitive group were instructed: “Please select left or right to make the generated sequence unpredictable.” Participants in the Random group were instructed: “Please select left or right to make the generated sequence as random as possible.”

In the main session, participants performed the sequence generation task (Fig. 1A). Each trial began with a cue, and participants had 1500 ms to make a binary choice by pressing a key for left or right. If a response was not made within this time frame, a warning message was displayed, and the trial was repeated. After the choice was made, their selected option was visually presented on the screen for 1500 ms. This feedback appeared exactly 1500 ms after the initial trial cue, regardless of the participant’s reaction time. The task consisted of 10 blocks, with each block comprising 20 trials. At the end of each block, participants rated on a 10-point scale (from 1 to 10) the extent to which they had kept their group’s specific instruction in mind. The entire experiment, including instructions and practice, lasted approximately 20 minutes.

Stimulus Generation for Experiment 2

The stimulus sequences for the prediction task in Experiment 2 were designed to capture the quintessential statistical properties of the human-generated sequences from Experiment 1. To avoid using an idiosyncratic sequence from a single participant, we generated two empirically calibrated sequences based on the aggregate behavior of the Competitive and Random groups.

For each condition, we constructed a generative model where the probability of repeating a choice was contingent on the current run length. The specific conditional probabilities for each run length were set to the group-average probabilities empirically determined in Experiment 1. We then ran a large-scale simulation (1 billion iterations) to generate candidate sequences, each consisting of 10 blocks of 20 trials. From this large pool, we selected the single sequence for each condition whose own conditional probability profile most closely matched the target group-average data. This process resulted in two final 200-trial stimulus sequences. For each participant in Experiment 2, the order of the 10 blocks within their assigned sequence was randomized.

Experiment 2: Sequence Prediction Task

Participants first received instructions on how to perform the task. They then completed a training session of 10 prediction trials to become familiar with the task, during which the correct answers were determined by a truly random process. In the main session, they performed the sequence prediction task (Fig. 2A). In each trial, participants had 1500 ms to predict the next item in a sequence by pressing a key for left or right. If a response was not made within this time frame, a warning message was displayed, and the trial was repeated. Following their response, both their choice and the correct choice were visually displayed for 1500 ms. This feedback was presented exactly 1500 ms after the initial trial cue, irrespective of their reaction time. The task consisted of 10 blocks of 20 trials each. Participants were not given any information about the origin or generative mechanism of the sequences (e.g., how or by whom they were created). At the end of each block, they rated their level of engagement with the prediction task on a 10-point scale. Experiment 2 lasted approximately 20 minutes.

Analysis

In Experiment 1, the primary dependent variables were the first-order and conditional probabilities of choice repetition. The conditional probability was calculated for each participant for runs of one to four identical preceding choices; longer runs were excluded from the analysis as they were too infrequent in the Random group to allow for robust statistical comparison. We used two-tailed one-sample t-tests to compare these probabilities against the 0.5 chance level and two-tailed independent-samples t-tests to compare the Competitive and Random groups. Additionally, to check for differences in how well participants kept their instructions in mind, we submitted their end-of-block ratings to a 2 (group: Competitive, Random) × 10 (block: 1–10) mixed-design ANOVA. To rule out potential confounds from simple response preferences, we also analyzed overall choice bias. We calculated both directional bias (the number of right choices minus the number of left choices) and the magnitude of bias (the absolute difference between them). Directional bias was compared between groups using an independent-samples t-test, while the magnitude of bias was compared using a non-parametric Wilcoxon rank-sum test.

We quantified the information content of the two empirically calibrated sequences (Experiment 2) in bits using standard Shannon measures. Let $X_t \in \{L\ (Left), R\ (Right)\}$ denote the correct item on trial *t*.

- Symbol entropy.

$$H(symbol) = -\sum_{X\in\{L,R\}} p(x)\log_2 p(x)$$

using the marginal frequencies of L/R across all trials.

- First-order (Markov) conditional entropy.
  We computed the 2×2 transition matrix $P(X_t|X_{t-1})$ within blocks (block boundaries excluded), the empirical distribution of the previous symbol $\pi(x)$, and
  $$H(X_t|X_{t-1}) = \sum_x \pi(x) H(P(\cdot\,|x))$$
  where $H(P(\cdot\,|x))$ is the binary entropy of the x-th row of P.
- Repetition/alternation entropy.
  Defining $R_t = 1\{X_t = X_{t-1}\}$, we report
  $$H(rep/alt) = -\sum_{r\in\{0,1\}} p(R_t = r)\log_2 p(R_t = r).$$
  This summarizes uncertainty about same vs different from the previous item.
- Run-length–conditioned entropies.
  For each run length ℓ (count of identical consecutive items at $t-1$), we computed
  $$H(X_t|run = \ell) \text{ and } H(rep/alt|run = \ell),$$
  and then formed frequency-weighted averages across ℓ. Rare long runs ($run \geq 6$) were pooled into a single bin.
- A coarse index of higher-order information beyond first order.
  As a single-number summary, we report
  $$\Delta H(run|prev) = H(X_t|X_{t-1}) - H(rep/alt|run)$$
  Larger ΔH indicates a greater reduction of uncertainty when run-length information is available. All entropies are in bits; transitions crossing block boundaries were excluded; long runs ($run \geq 6$) were pooled.

In Experiment 2, our main dependent variables were overall prediction accuracy (success rate) and the participants' expectation of repetition. We defined expectation of repetition as the probability that a participant's choice on a given trial matched the correct answer from the previous trial. We analyzed this expectation of repetition (1) overall across the experiment, (2) conditioned on the run length of the correct answer (from one to five), and (3) on a block-by-block basis (from block 1 to 10). To assess model calibration, we also calculated the difference between participants' conditional expectations and the true conditional probabilities of the empirically calibrated sequences. Statistical comparisons for these measures were performed using two-tailed one-sample t-tests (against the 0.5 chance level, or against a difference of zero for the

calibration analysis) and two-tailed independent-samples t-tests (between the two prediction groups), as appropriate. In addition to the block-wise analyses, we conducted a complementary 2 (Sequence: Competitive vs. Random; between-subjects) × 2 (Block: Block 1 vs. the mean of Blocks 2–10; within-subjects) mixed ANOVA on the expectation-of-repetition measure to summarize the early-to-late change. Similarly, to check for differences in task engagement, we submitted participants' end-of-block ratings to a 2 (group: predicting Competitive, predicting Random) × 10 (block: 1–10) mixed-design ANOVA.

To investigate the learning mechanisms underlying these prediction strategies, we conducted a trial-by-trial analysis of strategy updating. We defined two strategies on trial *t*: a "repetition strategy" if the participant's choice matched the correct answer from trial *t-1*, and an "alternation strategy" if it did not. The dependent variable was the probability of staying with the same strategy on trial *t+1*. This was submitted to a 2 (Sequence Type: Competitive vs. Random) × 2 (Previous Strategy on trial *t*: Repetition vs. Alternation) × 2 (Previous Outcome on trial *t*: Correct vs. Incorrect) mixed-design ANOVA.

For the model-based analysis, we specified a binary choice model for the probability of making a repetition prediction on trial *t*:

$$P_{rep}(t) = \sigma(\beta \cdot \left(Q_{rep}(t) - Q_{alt}(t)\right) + b_0 + \rho \cdot prevRep(t))$$

where $\sigma(x) = 1/(1 + e^{-x})$. where $Q_{rep}(t)$ and $Q_{alt}(t)$ denote the current values for the "predict repetition" and "predict alternation" strategies, respectively; β controls choice sensitivity to the value difference; b_0 is a baseline response bias (intercept); $prevRep(t) \in \{0,1\}$ indicates whether the previous strategy was a repetition (1) or an alternation (0); and ρ captures recency/inertia in repeating the previous strategy.

After feedback on trial *t*, we updated only the value of the strategy that was used, via separate learning rates for repetition and alternation:

$Q_{rep}(t + 1) = Q_{rep}(t) + \alpha_{rep}\left(r_t - Q_{rep}(t)\right)$ if the repetition strategy was chosen,

$Q_{alt}(t + 1) = Q_{alt}(t) + \alpha_{alt}\left(r_t - Q_{alt}(t)\right)$ if the alternation strategy was chosen,

with $r_t \in \{0,1\}$ indicating success (1) or failure (0) on trial *t*. Values were initialized at 0 and were not reset across blocks within each participant; values were reset across participants. The participant-wise parameter vector was

$$\theta = (\alpha_{rep}, \alpha_{alt}, \beta, b_0, \rho).$$

Parameters were estimated separately for each participant by minimizing the negative log-likelihood using constrained interior-point optimization (*fmincon.m*, multi-start).

Bounds were $\alpha_{rep} \in [-1,1], \alpha_{alt} \in [-1,1], \beta \in [0,20], b_0 \in [-5,5], \rho \in [-5,5]$; optimization used an optimality tolerance of 10^{-6} and up to 1,500 iterations with multiple distinct starting points.

Our a priori directional test evaluated whether the learning rate after choosing to repeat exceeds the learning rate after choosing to alternate; we applied a one-sided Wilcoxon signed-rank test to the participant-wise paired estimates and repeated the same directional test within each sequence context (Competitive vs. Random). To assess generalization across contexts, we compared the absolute level of each updating speed between contexts using two-sided rank-sum tests. For completeness, we tested whether the baseline response bias (intercept) and the recency/inertia term differed from zero in each sequence (one-sample signed-rank) and whether they differed between sequences (rank-sum).

All statistical analyses were conducted in MATLAB. For all sequential analyses, choice history was reset at the beginning of each block. For all t-tests, we report two-tailed p-values and Cohen's d as a measure of effect size. The significance threshold was set at $\alpha = 0.05$. We report exact p-values and effect sizes (r) for non-parametric tests throughout, reporting the standardized Z statistic for Wilcoxon signed-rank tests and the U statistic for Wilcoxon rank-sum tests. For ANOVAs, we report generalized eta-squared (η_G^2) as the measure of effect size. When the assumption of sphericity was violated, degrees of freedom were corrected using the Greenhouse-Geisser method. P-values for multiple comparisons within a family of tests were adjusted using the False Discovery Rate (FDR) correction method (Benjamini & Hochberg, 1995).

References

1. Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior research methods*, *53*(4), 1407-1425.
2. Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior research methods*, *52*(1), 388-407.
3. Armstrong, B. C., Frost, R., & Christiansen, M. H. (2017). The long road of statistical learning research: past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160047.
4. Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature neuroscience*, *7*(4), 404-410.

5. Barron, G., & Leider, S. (2010). The role of experience in the Gambler's Fallacy. *Journal of Behavioral Decision Making*, *23*(1), 117-129.
6. Beard, T. R., & Beil, R. O. (1994). Do people rely on the self-interested maximization of others? An experimental test. *Management Science*, *40*(2), 252-262.
7. Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations 1*, *2*, 69-186.
8. Benjamin, D. J., Moore, D. A., & Rabin, M. (2017). *Biased beliefs about random samples: Evidence from two integrated experiments* (No. w23927). National Bureau of Economic Research.
9. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289-300.
10. Bosch-Domenech, A., Montalvo, J. G., Nagel, R., & Satorra, A. (2002). One, two,(three), infinity,…: Newspaper and lab beauty-contest experiments. *American Economic Review*, *92*(5), 1687-1701.
11. Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. *The quarterly journal of economics*, *119*(3), 861-898.
12. Castillo, L., León-Villagrá, P., Chater, N., & Sanborn, A. (2024). Explaining the flaws in human random generation as local sampling with momentum. *PLOS Computational Biology*, 20(1), e1011739.
13. Chen, D. L., Moskowitz, T. J., & Shue, K. (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics*, *131*(3), 1181-1242.
14. Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, *36*(3), 181-204.
15. Costa-Gomes, M. A., & Crawford, V. P. (2006). Cognition and behavior in two-person guessing games: An experimental study. *American economic review*, *96*(5), 1737-1768.
16. Croson, R., & Sundali, J. (2005). The gambler's fallacy and the hot hand: Empirical data from casinos. *Journal of risk and uncertainty*, *30*(3), 195-209.
17. Fang, C., Aronov, D., Abbott, L. F., & Mackevicius, E. L. (2023). Neural learning rules for generating flexible predictions and computing the successor representation. *elife*, *12*, e80680.
18. Farmer, G. D., Warren, P. A., & Hahn, U. (2017). Who "believes" in the Gambler's

Fallacy and why?. *Journal of Experimental Psychology: General*, *146*(1), 63.

19. Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175-191.
20. Foucault, C., & Meyniel, F. (2021). Gated recurrence enables simple and accurate sequence prediction in stochastic, changing, and structured environments. *Elife*, *10*, e71801.
21. Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature reviews neuroscience*, *11*(2), 127-138.
22. Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical transactions of the Royal Society B: Biological sciences*, *364*(1521), 1211-1221.
23. Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in cognitive sciences*, *19*(3), 117-125.
24. Goeree, J. K., & Holt, C. A. (1999). Stochastic game theory: For playing games, not just for doing theory. *Proceedings of the National Academy of sciences*, *96*(19), 10564-10567.
25. Guseva, M., Bogler, C., Allefeld, C., & Haynes, J. D. (2023). Instruction effects on randomness in sequence generation. *Frontiers in Psychology*, *14*, 1113654.
26. Ho, T. H., Camerer, C., & Weigelt, K. (1998). Iterated dominance and iterated best response in experimental" p-beauty contests". *The American Economic Review*, *88*(4), 947-969.
27. Holt, C. A., & Roth, A. E. (2004). The Nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences*, *101*(12), 3999-4002.
28. Huang, Y., & Rao, R. P. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 580-593.
29. Jin, Y. (2021). Does level-k behavior imply level-k thinking?. *Experimental Economics*, *24*(1), 330-353.
30. Jokar, E., & Mikaili, M. (2012). Assessment of human random number generation for biometric verification. *Journal of Medical Signals & Sensors*, *2*(2), 82-87.
31. Kreps, D. M. (2018). Nash equilibrium. In *The new Palgrave dictionary of economics* (pp. 9251-9258). Palgrave Macmillan, London.
32. Lee, D., & Seo, H. (2011). Behavioral and neural variability related to stochastic choices during a mixed-strategy game. *The Dynamic Brain: An Exploration of*

Neuronal Variability and Its Functional Significance, 255-276.

33. Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *elife*, *8*, e41541.
34. McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica: Journal of the Econometric Society*, 803-836.
35. Misirlisoy, E., & Haggard, P. (2014). Asymmetric predictability and cognitive competition in football penalty shootouts. *Current Biology*, *24*(16), 1918-1922.
36. Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American economic review*, *85*(5), 1313-1326.
37. Nash Jr, J. F. (1950). Equilibrium points in n-person games. Proceedings of the national academy of sciences, 36(1), 48-49.
38. Nash, J. (1951). Non-Cooperative Games. *Annals of Mathematics*, *54*(2), 286–295.
39. Ochs, J. (1995). Games with unique, mixed strategy equilibria: An experimental study. *Games and Economic Behavior*, *10*(1), 202-217.
40. Oskarsson, A. T., Van Boven, L., McClelland, G. H., & Hastie, R. (2009). What's next? Judging sequences of binary events. *Psychological bulletin*, *135*(2), 262.
41. Rabin, M. (2002). Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, *117*(3), 775-816.
42. Rabin, M., & Vayanos, D. (2010). The gambler's and hot-hand fallacies: Theory and applications. *The Review of Economic Studies*, *77*(2), 730-778.
43. Rao, K., & Hastie, R. (2023). Predicting outcomes in a sequence of binary events: belief updating and gambler's fallacy reasoning. *Cognitive Science*, *47*(1), e13211.
44. Rapoport, A., & Budescu, D. V. (1997). Randomization in individual choice behavior. *Psychological Review*, *104*(3), 603.
45. Remillard, G. (2008). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *Quarterly Journal of Experimental Psychology*, *61*(3), 400-424.
46. Remillard, G. (2010). Implicit learning of fifth-and sixth-order sequential probabilities. *Memory & Cognition*, *38*(7), 905-915.
47. Remillard, G., & Clark, J. M. (2001). Implicit learning of first-, second-, and third-order transition probabilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 483.
48. Reny, P. J. (1999). On the existence of pure and mixed strategy Nash equilibria in discontinuous games. *Econometrica*, *67*(5), 1029-1056.
49. Roney, C. J., & Sansone, N. (2015). Explaining the gambler's fallacy: Testing a gestalt explanation versus the “law of small numbers”. *Thinking &*

Reasoning, *21*(2), 193-205.

50. Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, *8*(1), 164-212.
51. Santolin, C., & Saffran, J. R. (2018). Constraints on statistical learning across species. *Trends in Cognitive Sciences*, *22*(1), 52-63.
52. Schulz, M. A., Schmalbach, B., Brugger, P., & Witt, K. (2012). Analysing humanly generated random number sequences: a pattern-based approach. *PloS one*, *7*(7), e41531.
53. Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and cognition*, *112*, 92-97.
54. Stahl II, D. O., & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of economic behavior & organization*, *25*(3), 309-327.
55. Stahl, D. O., & Wilson, P. W. (1995). On players′ models of other players: Theory and experimental evidence. *Games and Economic Behavior*, *10*(1), 218-254.
56. Suetens, S., Galbo-Jørgensen, C. B., & Tyran, J. R. (2016). Predicting lotto numbers: a natural experiment on the gambler's fallacy and the hot-hand fallacy. *Journal of the European Economic Association*, *14*(3), 584-607.
57. Tishby, N., & Polani, D. (2010). Information theory of decisions and actions. In *Perception-action cycle: Models, architectures, and hardware* (pp. 601-636). New York, NY: Springer New York.
58. Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, *76*(2), 105.
59. Warren, P. A., Gostoli, U., Farmer, G. D., El-Deredy, W., & Hahn, U. (2018). A re-examination of “bias” in human randomness perception. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(5), 663.