Posterior parietal cortex mediates rarity-induced decision bias and learning under uncertainty

ABSTRACT

Making decisions when outcomes are uncertain requires accurate judgment of the probability of outcomes, yet such judgments are often inaccurate, owing to reliance on heuristics that introduce systematic errors like overweighting of low probabilities. Here, using a decision-making task in which the participants were unaware of outcome probabilities, we discovered that both humans and mice exhibit a rarity-induced decision bias (RIDB), *i.e.*, a preference towards rare rewards, which persists across task performance. Optogenetics experiments demonstrated that activity in the posterior parietal cortex (PPC) is required for the RIDB. Using *in vivo* electrophysiology, we found that rare rewards bidirectionally modulate choice-encoding PPC neurons to bias subsequent decisions towards rare rewards. Learning enhances stimulus-encoding of PPC neurons, which plays a causal role in stimulus-guided decisions. We then developed a dual-agent behavioural model that successfully recapitulates the decision-making and learning behaviours, and corroborates the specific functions of PPC neurons in mediating decision-making and learning. Thus, beyond expanding understanding of rare probability overweighting to a context where the outcome probability is unknown, and characterizing the neural basis for RIDB in the PPC, our study reveals an evolutionarily conserved heuristic that persistently impacts decision-making and learning under uncertainty.

INTRODUCTION

In a world that is probabilistic in nature, the ability to learn from uncertain outcomes and make optimal decisions is essential for survival[1-3]. This behaviour is an active research topic in economics, psychology, and neuroscience [1-6]. Decision making requires accurate judgment of event probabilities, yet these judgments are often inaccurate, as people tend to rely on heuristics and therefore make systematic errors when thinking under uncertainty [7, 8]. One well-established judgement error is the overweighting of low probabilities, *i.e.*, people tend to give more decision weights to rare events, and act as if they are more likely to happen than they really are [9-13]. This phenomenon results in a concave probability weighting function and constitutes a key component of prospect theory, a fundamental theory in economics [4, 10]. Studies have been done to explore the neural mechanisms underlying such probability distortion when outcome probabilities are explicitly shown to the subjects [11, 14, 15].

Previous work in humans and non-human primates has shown that the posterior parietal cortex (PPC) is activated by outcome uncertainty [16-20], and that PPC neurons represent the degree of uncertainty [21-23]. Studies in rodents have demonstrated a variety of functions of the PPC during decision-making: PPC neurons encode choices[24-28], visual stimuli [24, 26], and auditory stimulus categories [29]. Additionally, the PPC encodes short-term memory of stimulus, choice, and outcome [30-33], which are causally related to decision-making behaviour. PPC neurons have also been implicated in learning [26, 29].

Here, using an auditory stimulus-guided decision-making task with probabilistic outcomes, we demonstrate that when the outcome probability is unknown to the subjects, both humans and mice exhibit a rarity-induced decision bias (RIDB) affecting subsequent decisions. Our optogenetics experiments establish that PPC activity is required for RIDB, and *in vivo* recordings show that rare rewards bidirectionally modulate the activity of choice-selective PPC neurons to induce decision bias towards rarely rewarded choices. Moreover, as mice learned to make stimulus-guided decisions under uncertainty, PPC neurons were dynamically recruited for stimulus-guided decision-making. We then built a data-inspired dual-agent model that accurately reproduces the decision-making and learning behaviours observed for normal and PPC-inhibited conditions, and reveals essential impacts of PPC activity on the decision bias and learning.

RESULTS

Humans exhibit decision bias towards rarely rewarded choices

To assess the impact of rarity on decision-making under uncertainty, we designed a task with probabilistic outcomes (**Figure 1a**). Human subjects were instructed to make decisions based on auditory stimuli to collect maximal monetary rewards, without being informed of the probabilistic task structure. Pressing a start key triggers one of four auditory stimuli: duck, cat, canary, or frog sounds. After a duck or cat sound, pressing the "L" key leads to a monetary reward at 80% probability (common) and no reward at 20% probability (rare); pressing the "R" key leads to a reward of the same amount at 20% probability and no reward at 80% probability. Choice-reward associations were reversed for canary and frog sounds. All stimuli and common/rare outcomes are pseudorandomly interleaved (Methods), so the optimal strategy is to always make high-probability rewarded choices guided by the current stimulus, *i.e.*, always choose "L" for duck and cat, and "R" for canary and frog (termed stimulus-guided "correct" choices; **Supp. Fig. 1)**. The outcome was shown to the subjects as a money symbol or a no-money symbol upon each choice, but whether it was common or rare was not obviously shown.

Previous studies have established that uncertain outcomes have a strong influence on decision-making [2]. To quantitatively evaluate the impact of previous outcomes on future decision-making, we tracked the subjects' probability of repeating the preceding choices upon receiving 4 different outcomes: 80% rewarded (*comm*+), 20% rewarded (*rare*+), 80% unrewarded (*comm*-), and 20% unrewarded (*rare*-) (**Figure 1b**). We found that before subjects learned to reliably make stimulus-guided correct choices (some subjects never did; Methods, **Supp. Fig. 2a-b**), a rewarded choice leads to a higher choice repeating probability (RP) in the immediately next trial compared to an unrewarded choice (**Figure 1c**), indicating a win-stay, lose-switch (WSLS) tendency. To assess the impact of rare outcomes on future decision-making, we examined the difference in RP when the previous outcome was common vs. rare. Rare rewards triggered a slight increase in win-stay tendency, and unrewarded rare trials triggered a significant increase in lose-switch tendency (**Figure 1c**), indicating that rare outcomes have a stronger impact on subsequent decision-making than common outcomes.

Because this task is stimulus-guided by design, we also inspected the RP for the next same-stimulus trials (**Figure 1d**). We observed a much stronger WSLS tendency for the next same-stimulus trial than the immediately next trial (**Figure 1c, d**), suggesting that choices were more

strongly influenced by the preceding same-stimulus trial than the immediately preceding trial. Interestingly, we found a strong decision bias caused by rare rewards on the same-stimulus trials: subjects exhibited a significantly higher same-stimulus RP (sRP) following *rare*+ choices than *comm*+ choices (**Figure 1d**). We should emphasize that because *rare*+ choices are opposite to the correct choices, repeating *rare*+ choices resulted in much lower reward probability (=20%) than repeating *comm*+ choices (**Supp fig. 2c**). In spite of that, subjects still exhibited a significantly higher sRP_{rare+} than sRP_{comm+}, indicating a strong rarity-induced decision bias (RIDB) from preceding same-stimulus rewarded trials (**Figure 1d**). Thus, in an uncertain environment with probabilistic outcomes, rare rewards strongly influence subsequent decisions, revealing a rarity-induced heuristic that can contribute to overweighting of low probabilities under uncertainty.

A mouse model recapitulates the rarity-induced decision bias

To investigate the neural mechanisms underlying the RIDB observed in humans, we developed a behavioural paradigm in mice by adding probabilistic outcomes to a classical auditory two-alternative forced choice (2AFC) task [34] (**Figure 1e**). The mouse was allowed to freely poke into the centre port, triggering one of two auditory stimuli (A: 5k Hz pure tone; B: 13 kHz). Stimulus A was associated with 80% probability of water reward for left choices and 20% probability for right choices, and vice versa for stimulus B. All mice trained in this task showed a sigmoidal learning curve (performance for each session, calculated as the percentage of correct choices; **Supp Fig. 3a**), similar to an auditory discrimination task with certain outcomes [35].

Like humans, mice exhibited a WSLS tendency both in the immediately next and next same-stimulus trials (**Figure 1f**), before they learned to reliably make correct choices (Methods, **Supp Fig. 3**). Also, like humans, mice exhibited significantly higher sRP_{rare+} than sRP_{comm+} (**Figure 1g**). Therefore, our behavioural paradigm for mice successfully recapitulates the strong RIDB initially found in humans, pointing to an apparently conserved rarity-induced heuristic manifest as a tendency to repeat rarely rewarded choices despite suboptimal outcomes.

We next investigated whether the strength of the RIDB was associated with the rarity levels. We modified the common/rare trial ratio (90%:10%, n = 4 mice; 95%:5%, n = 5 mice, **Figure 1g**), and quantified ΔsRP $(sRP_{rare+} - sRP_{comm+})$. Briefly, we observed larger ΔsRP with elevated common/rare ratios (**Figure 1h**), indicating that the rarer the rewards, the stronger the rarity influence. Thus, beyond detecting a rarity-induced heuristic in both humans and mice, we

established an informative empirical system for studying the neural basis of this conserved heuristic.

RIDB persists across all performance levels

Another informative trend was that in mice as well as in humans, as the performance improves, ΔsRP gradually declines (**Figure 1i, j**). Although it is obvious that a positive ΔsRP value at low performance levels reveals the existence of RIDB, it is important to clarify that ΔsRP < 0 does not indicate the disappearance of RIDB. This is because, in addition to the influence of preceding trials, through which RIDB manifests, there also exists a stimulus-guided factor that impacts choices. This factor is related only to the current stimulus and thus independent of RIDB, yet modulating sRP_{comm+} and sRP_{rare+} in opposite directions: it pushes choices towards *comm*+ and away from *rare*+, thus increasing sRP_{comm+} and reducing sRP_{rare+}. As performance improves, such modulation becomes increasingly significant.

To assess RIDB regardless of performance levels, we define D_{rare+} and D_{comm+} as a measure of the influence of *rare*+ and *comm*+ on subsequent choices, respectively (Methods). Thus, RIDB can be evaluated as (D_{rare+} - D_{comm+}). To compute D_{rare+}, we measured the difference between the conditional probability of making a *rare*+ choice following a *rare*+ trial, and the probability of making a *rare*+ choice in the same session following any trial:

$$D_{rare+} = \mathrm{p}\ (\mathrm{rare+|preceding\ rare+}) - \mathrm{p(rare+)}$$

Note that p (rare+|preceding rare+) is the definition of sRP_{rare+}, and p(rare+) equals to (1 – performance), performance being the averaged performance of the session. Similarly,

$$D_{comm+} = \mathrm{p}\ (\mathrm{comm+|preceding\ comm+}) - \mathrm{p(comm+)}$$

Here, p (comm+|preceding comm+) is sRP_{comm+}, and p(comm+) equals to performance.

Results show that D_{rare+} is consistently larger than D_{comm+} across all performance levels for humans and mice (**Figure 1k, l**), indicating that RIDB persisted throughout task performance. Thus, rare rewards have a strong impact on decision-making even after animals were extensively exposed to the task with outcome uncertainty and had learned to predominantly make stimulus-guided decisions.

PPC activity is necessary for generating RIDB in mice

Previous studies in rodents have shown that the posterior parietal cortex (PPC) encodes behavioural history [30-33], which is ostensibly related to the RIDB. To investigate whether the PPC is involved in mediating RIDB, we used optogenetics (eNpHR3.0) to bilaterally inhibit the PPC in rare trials, and compared sRP_{rare+} with sRP_{comm+} in the PPC-inhibited mice (**Figure 2a, Supp. Fig. 4**). If PPC activity is required for RIDB, then we would expect sRP_{rare+} to be similar to sRP_{comm+} in the PPC-inhibited mice.

We first inhibited the PPC during the outcome epoch of the rare trials, as this is the epoch during which mice apparently perceive rarity (n=5, *outcome-opto*, see Methods). A masking yellow light was delivered in all trials to ensure that mice did not treat the stimulating light as an additional cue (Methods). To our disappointment, sRP_{rare+} is still higher than sRP_{comm+} in *outcome-opto* mice (**Figure 2b, c**). However, for a separate group of mice in which we inhibited the PPC during the stimulus epoch of rare trials (n=5, *stim-opto*), we found a significant reduction in ΔsRP as compared to non-inhibited controls (**Figure 2b, c**), across a wide range of performance levels (**Figure 2d, e, Supp. Fig. 4**), corroborating the notion that RIDB persisted even at high task performance. These results indicate that PPC activity is required for the RIDB.

Rare rewards bidirectionally modulate choice-selective PPC neurons to induce RIDB

To investigate the neural mechanisms underlying the RIDB in the PPC, we made extracellular recordings in the mouse PPC during task performance (5 mice, 100 behavioural sessions, 3135 units, **Figure 3a, Supp. Fig. 5, 6a**). Previous rodent studies have shown that PPC activity can drive animal's choice [12, 19, 20]. Thus, it is plausible that preceding *rare*+ trials somehow modulate PPC activity to influence subsequent choices. To investigate the rarity-induced modulation, we first examined how a preceding *rare*+ trial influences PPC neuronal activity in the next same-stimulus trial. We separated trials into 4 groups based on the outcome of the preceding trial (**Figure 3a, b**). Then, we used ROC analysis to calculate a rarity modulation index (RMI) that measures how strongly a neuron's firing rates following *rare*+ trials differ from that neuron's firing rates following other trials. We observed that ~15% of recorded PPC neurons were differentially modulated by preceding *rare*+ trials compared to other trials (**Figure 3c**), among which ~half were significantly enhanced by preceding rare+ (RMI < 0, **Figure 3c**) and ~half were significantly suppressed (RMI > 0, **Figure 3c**). These results establish that *rare*+ modulates PPC neuronal activity in the next same-stimulus trial. Further, and consistent with the aforementioned persistence

of RIDB across performance levels, we found that PPC activity was modulated by preceding *rare*+ trials at all performance levels (**Supp. Fig. 7a, b**).

We subsequently investigated how this rarity-induced modulation of PPC activity influences choices (**Figure 3d**). Using ROC analysis to compute the choice preference index of individual PPC neurons, we found that a majority of PPC neurons either preferred right choices (*i.e.,* firing rates higher for right than left choices) or left choices (**Supp. Fig. 6**). Upon combining the choice preference index with the RMI for each neuron, we found that for stimulus A, for which *rare*+ choices are right choices (**Figure 3d**), most neurons that were enhanced by preceding *rare*+ preferred right choices (**Figure 3e, f**). In contrast, most neurons that were suppressed by preceding *rare*+ preferred left choices (**Figure 3e, f**). These trends are both consistent with the right bias induced by stimulus A *rare*+. For stimulus B, for which *rare*+ choices are left choices, preceding *rare*+ enhanced left-preferred neurons and suppressed right-preferred neurons (**Figure 3d-f**), consistent with a left bias. Again, such rarity-induced modulation biased choices towards *rare*+ choices across all performance levels **(Supp. Fig. 7)**. These results establish a neural basis for how *rare*+ trials influence subsequent decisions (*i.e.*, a neural basis for RIDB): by bi-directionally modulating choice-encoding PPC neurons to induce bias towards *rare*+ choices.

Learning under uncertainty enhances stimulus-encoding of PPC neurons to drive decisions

We demonstrated earlier that most humans and all mice eventually learned to make stimulus-guided correct decisions. It has been established that the PPC participates in learning under uncertainty [19, 36], but whether PPC activity is required for learning under uncertainty was unclear. In *stim-opto* mice, in which we inhibited PPC during the stimulus epoch of 20% of trials, we observed an increasing impairment on performance of PPC-inhibited trials as the mouse's performance improves (**Figure 4a, b**). The impairment became significant when mice reached high performance (>85%, termed "expert stage", **Figure 4c, d**). No difference was observed for *outcome-opto* mice during learning (**Figure 4a)**, but after these mice reached expert stage, as soon as we changed the inhibition window to the stimulus epoch (*stim-opto acute*), they immediately exhibited impaired performance in PPC-inhibited trials (**Figure 4a, c-d**). These findings demonstrate that the PPC activity during the stimulus epoch plays a causal role in stimulus-guided decision-making.

To analyse how PPC neuronal activity evolves during learning, we conducted a two-way ANOVA considering two factors: stimulus and choice [24-28]. Whereas the proportion of choice-encoding neurons remained consistently high throughout learning (**Figure 4e**), the proportion of stimulus-encoding neurons varied significantly: it remained low before performance reached 65%, rose quickly paralleling sharp increase of the performance, and stayed high until performance reached 85% (**Figure 4f**). These results suggest that stimulus-guided learning under uncertainty requires a dynamic recruitment of PPC neurons, especially during the quick rise of stimulus dependency in decision-making.

The rise followed by a decline of stimulus-encoding neurons (**Figure 4e**) seems to contradict the monotonically increasing impairment caused by PPC inhibition as performance improves (**Figure 4b**). A possible explanation is that although the number of stimulus-encoding neurons fluctuates, the magnitude of stimulus-encoding increases with learning. To test this, we computed the stimulus discriminability (d') by comparing the firing rates of each neuron's responses to stimulus A vs. B [37] (**Figure 4f**), and calculated the AUC (area under curve) of d' during the stimulus epoch (**Figure 4g**). We found that the stimulus d' of individual PPC neurons increased monotonically during learning (**Figure 4f, g**). We then used demixed PCA [38] to assess the variance explained by the stimulus at the populational level (**Supp Fig. 7**). This analysis demonstrated an extremely low variance explained by the stimulus at low performance levels, followed by a continuous increase (**Figure 4h**). Thus, as learning progresses, an increasing magnitude of stimulus-encoding by the PPC neurons underlies their causal impact on stimulus-guided decision-making.

A data-inspired behavioural model reveals impact of PPC activity on learning speed

We have demonstrated a persistent RIDB across performance levels mediated by the PPC (**Figure 2d, e**), and an increasing stimulus-dependency of decision-making encoded by PPC neurons as learning progresses (**Figure 4b**), but it remains unclear how RIDB interacts with learning. To understand the potential interaction, we built a dual-agent model. The first agent is a stimulus-specific, rarity-state-based win-stay-lose-switch (rs-WSLS) agent [39] (**Figure 5a**). The rs-WSLS rule is based on the finding that mice exhibited a strong stimulus-specific WSLS tendency **(Figure 1f)**, and the existence of rarity state throughout learning is inspired by the behavioural and neural data showing the persistent RIDB (**Figure 2d, e; Supp Fig. 2d**). In this model, the win-stay and

lose-switch probabilities are updated trial-by-trial, depending on the internal estimate of the rarity state (Methods) [1, 40, 41], which can effectively track the occurrence of rare trials (**Supp Fig. 8**).

The second agent is a stimulus-guided reinforcement learning (RL) model [42] (**Figure 5a**). The RL agent uses a multiplicative update rule [43] and different learning rates for rewarded and unrewarded trials [44]. In light of the behavioural and neural data demonstrating the persistent RIDB and gradually increasing stimulus-dependency, we used normalized decision weights derived from entropy-based metrics in the RL model to combine the two agents [45] (Methods). We found that the integrated model successfully recapitulated the RIDB we observed at 20%, 10% and 5% rarity levels (**Figure 5b**), as well as the relationship between ΔsRP and performance (**Figure 5c**). This model also replicated two phenomena observed in the optogenetic inhibition of the PPC during the stimulus epoch: 1, throughout learning, invalidating the rarity state in the model reduced ΔsRP (**Figure 5d**), reproducing the observed elimination of RIDB by PPC inhibition (**Figure 2d**); 2, in the expert stage, invalidating the RL impaired performance (**Figure 5e**), reproducing the observed impairment of stimulus-guided decisions (**Figure 4c-d**).

Our model predicts faster learning speeds for lower rarity levels, despite the larger RIDB, which closely resembles the learning speeds we observed for 20%, 10%, and 5% rarity levels in mice (**Figure 5f**). Our model also predicts that learning would be accelerated by inhibiting PPC activity during rare trials (**Figure 5g**), which was indeed observed in *stim-opto* mice (**Figure 5g**). Thus, our model supports the essential influence of PPC activity on learning, as validated by optogenetic experiments.

Discussion

To navigate in a world where uncertainty is universal, one needs to make good assessments of the probabilities of upcoming events. Here, using a behavioural paradigm with probabilistic outcomes, we show that in both humans and mice, rarity causes a persistent decision bias towards rare rewards. We demonstrate that the PPC is required for generating this decision bias, and show that PPC neurons are modulated by preceding rare rewards to induce the decision bias. Upon tracking PPC neuronal activity throughout learning, we observed a dynamic change in stimulus-encoding of PPC neurons (in both the number of neurons and the magnitude of encoding). This stimulus-encoding by the PPC is behaviourally relevant, as PPC inhibition increasingly deteriorates performance as learning progresses. Finally, we built a dual-agent behavioural model which successfully predicted that PPC inhibition during rare trials would accelerate learning, further demonstrating the causal influence of the PPC on decision-making and learning under uncertainty.

The ability to adjust one's behaviour in response to uncertain outcomes is essential for survival, and it is well-established that the subjective assessment of probability in uncertain environments is often biased and results in systematic errors [8]. These errors are caused by many factors such as risk-seeking for rewards, risk-aversion for losses, ease of remembering familiar or extreme events, and emotions [4, 10]. In the present study, we found that in a task where both humans and mice were motivated to obtain maximal rewards, they exhibited a systematic decision bias: preference for rarely rewarded over commonly rewarded choices. As this task excludes influences from risk, reward size, and emotions, our results support that rarity *per se* induces a heuristic that causes humans and mice to inaccurately estimate the reward probability associated with their choices. We also show that this decision bias persists even when the subjects have been extensively exposed to the probabilistic outcomes and are capable of performing the task at an expert level, findings consistent with the notion that "the reliance on heuristics and the prevalence of biases are not restricted to laymen"[8]. Further, this decision bias aligns with commonly reported overweighting of small probabilities in studies related to prospect theory, in which the outcome probabilities are explicit and evaluated consciously by the subjects [4, 9, 10, 12]. Our study has thus uncovered a conserved, covert preference for rare rewards, extending this well-known phenomenon to a condition in which the probability is not obviously revealed, a common real-world scenario [46] .

Previous studies have speculated that animals adopt different learning strategies depending on the uncertainty type: fast learning under unexpected uncertainty (changing environment) and slow learning under expected uncertainty (stable environment) [2, 47, 48]. Our dual-agent behavioural model complies with this hypothesis: 1) the fast-updating rs-WSLS agent represents RIDB in decisions, coinciding with the unexpected uncertainty in the early phase of the task, as subjects were not informed of the outcome probabilities *a priori*; 2) the slow-updating RL agent represents stimulus-guided learning, achieved by the subjects only when they figured out the outcome probabilities after extensive task exposure, consistent with slow learning under expected uncertainty. Our data indicate that the PPC is able to represent both fast-updating rs-WSLS and slow-updating RL agents, and it is possible that the neuronal populations representing these agents communicate with distinct brain regions and neuromodulatory systems[16, 29, 31, 49]. Future studies are needed to further understand the neural circuit mechanisms.

Many studies in primates and rodents have established that the PPC is involved in decision-making [26-28, 50, 51]. Specifically, the PPC has been linked to decision-making under uncertainty [21-23]. Notably, the studies concerning uncertainty were mostly conducted in primates and did not include manipulation experiments probing the causal role of the PPC in the decision process. Our experiments demonstrate two separate roles for the PPC in decision making under uncertainty: 1) mediating the decision bias towards rare rewards, and 2) mediating the stimulus-guided decisions at the expert stage. Our study has thus filled in a gap in understanding about how the PPC is causally involved in learning under uncertainty, specifically demonstrating that PPC activity is dynamically modulated by learning and, moreover, is causally related to auditory stimulus-guided decision-making as learning progresses.

Figure captions

Figure 1. Rarity induces a persistent decision bias towards rarely rewarded choices in humans and mice

(a) Schematic showing an auditory decision-making task with probabilistic outcomes for humans. Pressing "S" key triggers the delivery of 1 of 4 animal sounds (duck, cat, canary or frog). Upon hearing duck or cat sounds, pressing "L" key leads to money reward at 80% probability, and no money at 20% probability, and pressing "R" key leads to money reward at 20% probability, and no money at 80% probability. The choice-reward contingency is opposite for canary and frog sounds.

(b) Schematic showing choice repeating for the immediately next trial (in which the auditory stimulus can be the same as or different from the preceding trial) and the next same-stimulus trial (which may be directly following the preceding trial or with other stimuli in between) after each possible outcome: reward in a common trial (*comm*+), no reward in a rare trial (*rare*-), no reward in a common trial (*comm*-), and a reward in a rare trial (*rare*+).

(c) Left panel, probabilities of repeating a rewarded (blue) and an unrewarded (red) choice in the immediately next trial (nRP). Middle panel, nRP for *comm*+, *rare*+, *comm*-, or *rare*- trials. Right panel, difference in the probability (ΔRP) of repeating a rewarded vs. an unrewarded trial (blue), and ΔRP of repeating a *rare*+ vs. a *comm*+ trial (green). Wilcoxon signed-rank test.

(d) Same as (c), for next same-stimulus trials.

(e) Schematic showing an auditory 2-alternative-forced choice (2AFC) task with probabilistic outcomes for mice. Poking in the centre port triggers the stimulus A (5kHz pure tone) or B (13kHz pure tone). For stimulus A, left choice is associated with water reward at 80% probability and timeout punishment at 20%; right choice is associated with timeout at 80% probability and water at 20%. The choice-reward contingency is the opposite for stimulus B.

(f) Left panel, probabilities of repeating a rewarded (blue) and an unrewarded (red) choice in the immediately next trial. Right panel, probabilities of repeating a rewarded and an unrewarded choice in the next same-stimulus trial. Wilcoxon signed-rank test.

(g) Probabilities of repeating a *comm*+, *rare*+, *comm*-, or *rare*- choice in the next same-stimulus trial for 3 rarity levels: 20%, 10%, and 5%. Wilcoxon signed-rank test.

(h) ΔsRP of repeating a *rare*+ vs. *comm*+ trial for 3 rarity levels in (g).

(i) ΔsRP computed for human subjects using trials cumulated over 40, 80, 120, 160 and 200 trials, against task performance.

(j) ΔsRP computed for every session of all mice as performance increases.

(k) Each behavioural session is divided into 40-trial blocks *post hoc* for analysis. Left, D_{rare+} and D_{comm+} for every 40-trial block of all human subjects, plotted against block performance. Right, D_{rare+} and D_{comm+} averaged across performance. The difference represents RIDB.

(l) Left, D_{rare+} and D_{comm+} for every session of all mice, plotted against session performance. Right, D_{rare+} and D_{comm+} averaged across performance. The difference represents RIDB.

Figure 2. Optogenetic inhibition of the PPC eliminates RIDB.

(a) Illustration showing optogenetic experiment. Optogenetic inhibition of PPC was performed either during stimulus epoch (*stim opto*) or outcome epoch (*outcome opto*) in rare trials.

(b) Bar graph showing sRP_{rare+} and sRP_{comm+} for *stim opto* (left, n=5) and *outcome opto* (right, n=5) mice. Paired t-test.

(c) Comparison of ΔsRP for *non-opto* (control), *outcome opto* and *stim opto* mice. Wilcoxon ranksum test.
(d) Correlation between ΔsRP and performance for *non-opto* (grey), *stim-opto* (red) and *outcome-opto* (orange) mice. Each data point represents a behavioural session.
(e) Comparison of correlation between ΔsRP and performance for *stim-opto* and *non-opto* mice.

Figure 3. Rare rewards modulate choice-selective PPC neurons to induce decision bias

(a) Top panel, illustration showing the tetrode recording setup. Bottom panel, schematic showing the modulation by a preceding *rare*+ trial on the PPC neuronal activity during the next same-stimulus trial.
(b) Left: top panel, raster plot showing spikes of an example PPC neuron enhanced by preceding *rare*+ trials; bottom panel, averaged firing rates of the example neuron, showing enhanced firing (green) by *rare*+. Right: top panel, raster plot showing an example PPC neuron suppressed by *rare*+; bottom panel, averaged firing rates of the example neuron, showing suppressed firing (dark green) by *rare*+.
(c) Histogram showing rarity modulation index (RMI) of PPC neurons by *rare*+ in stimulus A trials (top panel) and stimulus B trials (bottom panel). Neurons with mean firing rates of >2 Hz recorded in sessions with ≥3 *rare*+ trials were included in the analysis. For each neuron, RMI was calculated using ROC analysis. Statistical significance was assessed via bootstrapping (200 iterations, $p<0.01$).
(d) Top panel, schematic showing a preceding stimulus A *rare*+ trial (right choice, inducing a right bias) can enhance the firing of right-preferred PPC neurons and suppress the firing of left-preferred PPC neurons during the next same-stimulus trial. Bottom panel, schematic showing a preceding stimulus B *rare*+ trial (left choice, inducing a left bias) can enhance the firing of left-preferred PPC neurons and suppress the firing of right-preferred PPC neurons during the next same-stimulus trial.
(e) Top panel, example choice-selective neurons modulated by stimulus A *rare*+: right-preferred neuron (example 1) is enhanced; left-preferred neuron (example 2) is suppressed. Bottom panel, example choice-selective neurons modulated by stimulus B *rare*+: left-preferred neuron (example 3) is enhanced; right-preferred neuron (example 4) is suppressed.
(f) Rarity modulation index (RMI) of Stimulus A *rare*+ (top panel) and stimulus B *rare*+ (bottom panel) plotted against choice preference index for all neurons. Each data point represents one neuron. Red data points indicate neurons with significant choice preference and rarity modulation. Histograms show choice preference index distribution for *rare*+ enhanced (right bar graph) and suppressed (left bar graph) neurons.

Figure 4. causal role of the PPC in stimulus-guided decisions and evolvement of stimulus-encoding in PPC neurons

(a) Example performance of a *stim-opto* (left) and an *outcome-opto* (right) mouse, each showing averaged performance of non-opto and opto trials. The *outcome-opto* mice was then switched to *stim-opto* protocol for 3 sessions.
(b) Correlation between the reduction in performance of *stim-opto* trials compared to non-opto trials and the performance of non-opto trials. Each data point represents one behavioural session.
(c) Performance of non-opto and opto trials in *stim-opto (n=5)* and *outcome-opto (n=5)* mice, when non-opto trial performance of a session reaches 85% performance level (expert stage).

Stim-opto (acute) represents data from 4 *outcome-opto* mice that were switched to inhibition during stimulus epoch after reaching 85% performance. Paired t-test.

(d) Difference in performance between non-opto and opto trials in *stim-opto (n=5), outcome-opto (n=5),* and *Stim-opto (acute)* (n=4) mice.

(e) Fraction of significant choice-encoding (green) and stimulus-encoding (purple) neurons at different choice accuracies. Significance is determined using two-way ANOVA ($p<0.01$).

(f) Mean stimulus discriminability (d') for stimulus-encoding neurons recorded in sessions in which the mouse performance <65%, 65%<performance<75%, 75%<performance<85%, or performance>85%. Mean ± s.e.m.

(g) AUC (area under curve) of stimulus d' during the stimulus epoch of stimulus-encoding neurons. Each data point indicates one neuron.

(h) Demixed PCA (dPCA) analysis showing population stimulus encoding at low (performance <65%) and high (>85%) performance levels.

(i) dPCA analysis showing population stimulus coding across different performance levels.

Figure 5. A dual-agent behavioural model showing the impact of PPC on decision-making and learning

(a) Schematic showing the behavioural model, consisting of a rarity-based WSLS agent (top) and a RL agent (bottom).

(b) Simulated repeat probability for next same-stimulus trials at 20%, 10% and 5% rarity levels.

(c) Simulated ΔsRP across performance levels at 20%, 10% and 5% rarity levels.

(d) Simulated ΔsRP across performance levels at 20% rarity level, upon elimination of rarity state in the model.

(e) Simulated performance impairment at expert stage by inactivation of the RL agent.

(f) Trials to reach expert stage at 20%, 10% and 5% rarity levels, for model (left) and mouse (right).

(g) Trials to reach expert stage at 20% rarity level upon simulated inactivation of the PPC in the model (left), reproduced by optogenetic PPC inhibition in mice (right).

Supplementary Figure 1. Strategies to achieve maximum reward in this task.

(a) Example trials in a session, demonstrating a subject's choices and the corresponding outcomes received. The subject's choices were consistently biased by preceding *rare+*.

(b) Choice strategies and corresponding reward rates. To achieve maximal rewards, the optimal strategy is to always choose the *comm+* choices, yielding a maximal reward rate of 80%. Other choice strategies produce lower reward rates. The reward rates were produced by averaging the reward rates of 100-iteration simulations.

Supplementary Figure 2. Learning curves and repeat probabilities of all human subjects.

(a) The learning curves of all subjects that eventually learned to make stimulus-guided choices. Dashed line indicates a performance threshold of 85%, above which subjects are considered to have learned the task. Performance is evaluated with a running average of 20 trials.

(b) The learning curves of subjects that failed to learn the task.

(c) The relationship between theoretical reward rate and sRP_{rare+} and sRP_{comm+}.
(d) Rare outcomes did not affect the repeat probability of the immediate next trial, not considering different outcomes. Wilcoxon sign-rank test.

Supplementary Figure 3. Mice also exhibited the RIDB in next same-stimulus trials, but not in the immediately next trials.
(a) Segmentation of learning into naïve, intermediate and expert stages using a sigmoid fitting. Stages were segmented by 20% and 80% max performance in each animal.
(b) The fitted number of sessions in naïve or before expert stage.
(c) The learning curves of all animals in the 0.8:0.2 (common:rare) group.
(d) The learning curves of all animals in the 0.9:0.1 group.
(e) The learning curves of all animals in the 0.95:0.05 group.
(f) nRP for 0.8:0.2, 0.9:0.1, 0.95:0.05 groups.
(g) ΔsRP across different performance levels for 0.8:0.2, 0.9:0.1, 0.95:0.05 groups. Each data point represents one session. Each colour represents one mouse.

Supplementary Figure 4. Optogenetic inhibition of the PPC.
(a) Histology showing the optogenetic manipulation site. Left, side-by-side comparison with the Allen Atlas. Right, zoomed-in view of the virus expression. The yellow rectangle indicates the position of the optic fibre. Scale bar, 0.5mm.
(b) Effects of the optogenetic inhibition on nRP for *outcome-opto* and *stim-opto* mice. Two-sample paired t-test.
(c) Left, D_{rare+} and D_{comm+} for every session of *outcome-opto* mice, plotted against session performance. Right, D_{rare+} and D_{comm+} averaged across performance. The difference represents RIDB.
(d) Left, D_{rare+} and D_{comm+} for every session of *stim-opto* mice, plotted against session performance. Right, D_{rare+} and D_{comm+} averaged across performance. The difference represents reduced RIDB compared to control and *outcome-opto* mice.

Supplementary Figure 5. Behavioural performance of the mice in tetrode recording experiments.
(a) Learning curves of all 5 mice used in tetrode recording. Each colour represents one mouse.
(b) nRP and sRP for *comm+, rare+, comm-* and *rare-* trials of the recorded mice. Two-sample paired t-test.
(c) Comparison of ΔsRP between mice in recording experiments and previously reported control mice. Wilcoxon rank sum test.
(d) Relationship between ΔsRP and performance of recorded mice.

Supplementary Figure 6. *In-vivo* extracellular recording of the PPC neurons in mice.
(a) Histology showing the tetrode trace in the PPC. Tetrodes were dipped in DiI before implantation. Yellow arrow indicates the deepest location being recorded. Scale bar, 1 mm.
(b) Peri-stimulus time histogram (PSTH) of an example choice preferring PPC neuron.
(c) Histogram showing choice preference index (CPI) of PPC neurons. Neurons with mean firing rates of >2 Hz recorded in sessions with ≥3 rare+ trials were included in the analysis. For each neuron, CPI was calculated using ROC analysis. Statistical significance was assessed via bootstrapping (200 iterations, $p<0.01$).

(d) Relationship between the CPI and corresponding session performance. PPC neurons exhibited stable choice preference across performance levels.
(e) Fraction of significant ($p<0.01$) choice preferring neurons, at different choice performance levels. Each fraction value is assessed with a sliding window of ±2.5% performance.
(f) Averaged d' for choice selective neurons recorded in sessions with different performance levels.
(g) AUC of d' for choice selective neurons across task performance.

Supplementary Figure 7. Outcome modulation on PPC neurons.

(a) Relationship between RMI and performance for each recorded neuron. The magnitude of modulation (the absolute value of RMI) is correlated with performance. Note that positive RMI reflects *rare+* enhancement, and negative RMIs reflect *rare+* suppression of neuronal activity.
(b) Relationship between the percentage of neurons with significant modulation and corresponding sessions' performances.
(c) Intersection of rare+ modulated and choice-preferring neurons.
(d) The relationship between (*rare+&rare-*) modulation and CPI, showing modulation of PPC neurons by rare trials (combining both rewarded and unrewarded trials). The modulation index of (*rare+&rare-*) trials are estimated with balanced choices in the preceding same-stimulus and current trials, thus eliminating the potential influence of choice history. K-S test is performed on the distributions of modulated choice preferences.

Supplementary Figure 8. Rarity state detection in the model.
(a) Actual rare trials and rare trials estimated by the model, given the trial sequence.
(b) Correlation of actual and estimated rare trials.

References

1. Soltani, A. and A. Izquierdo, *Adaptive learning under expected and unexpected uncertainty.* Nature Reviews Neuroscience, 2019. **20**(10): p. 635-644.
2. Monosov, I.E., *How Outcome Uncertainty Mediates Attention, Learning, and Decision-Making.* Trends in Neurosciences, 2020. **43**(10): p. 795-809.
3. Platt, M.L. and S.A. Huettel, *Risky business: the neuroeconomics of decision making under uncertainty.* Nature Neuroscience, 2008. **11**(4): p. 398-403.
4. Tversky, A. and D. Kahneman, *Advances in prospect theory: Cumulative representation of uncertainty.* Journal of Risk and Uncertainty, 1992. **5**(4): p. 297-323.
5. Lee, D., H. Seo, and M.W. Jung, *Neural Basis of Reinforcement Learning and Decision Making.* Annual Review of Neuroscience, Vol 35, 2012. **35**: p. 287-308.
6. Fiser, J., et al., *Statistically optimal perception and learning: from behavior to neural representations.* Trends in Cognitive Sciences, 2010. **14**(3): p. 119-130.
7. Tversky, A. and D. Kahneman, *Availability: A heuristic for judging frequency and probability.* Cognitive Psychology, 1973. **5**(2): p. 207-232.

8. Tversky, A. and D. Kahneman, *Judgment under Uncertainty: Heuristics and Biases.* Science, 1974. **185**(4157): p. 1124-1131.
9. Elston, T.W., et al., *Extreme Outcomes Accentuate Overweighting of Low Probabilities and Underweighting of High Probabilities in Experience-Based Choice.* Journal of Behavioral Decision Making, 2024. **37**(2): p. e2380.
10. Kahneman, D. and A. Tversky, *Prospect Theory: An Analysis of Decision under Risk.* Econometrica, 1979. **47**(2): p. 263-291.
11. Constantinople, C.M., A.T. Piet, and C.D. Brody, *An Analysis of Decision under Risk in Rats.* Current Biology, 2019. **29**(12): p. 2066-+.
12. Stauffer, W.R., et al., *Economic Choices Reveal Probability Distortion in Macaque Monkeys.* Journal of Neuroscience, 2015. **35**(7): p. 3146-3154.
13. Hsu, M., et al., *Neural Response to Reward Anticipation under Risk Is Nonlinear in Probabilities.* Journal of Neuroscience, 2009. **29**(7): p. 2231-2237.
14. Tymula, A., et al., *Dynamic prospect theory: Two core decision theories coexist in the gambling behavior of monkeys and humans.* Science Advances, 2023. **9**(20).
15. Imaizumi, Y., et al., *A neuronal prospect theory model in the brain reward circuitry.* Nature Communications, 2022. **13**(1).
16. Lopez-Paniagua, D. and C.A. Seger, *Coding of level of ambiguity within neural systems mediating choice.* Frontiers in Neuroscience, 2013. **7**.
17. Huettel, S.A., A.W. Song, and G. McCarthy, *Decisions under uncertainty: Probabilistic context influences activation of prefrontal and parietal cortices.* Journal of Neuroscience, 2005. **25**(13): p. 3304-3311.
18. Krain, A.L., et al., *Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making.* Neuroimage, 2006. **32**(1): p. 477-484.
19. Bach, D.R., et al., *The Known Unknowns: Neural Representation of Second-Order Uncertainty, and Ambiguity.* Journal of Neuroscience, 2011. **31**(13): p. 4811-4820.
20. Huettel, S.A., et al., *Neural signatures of economic preferences for risk and ambiguity.* Neuron, 2006. **49**(5): p. 765-775.
21. Horan, M., N. Daddaoua, and J. Gottlieb, *Parietal neurons encode information sampling based on decision uncertainty.* Nature Neuroscience, 2019. **22**(8): p. 1327-+.
22. Yang, T. and M.N. Shadlen, *Probabilistic reasoning by neurons.* Nature, 2007. **447**(7148): p. 1075-U2.
23. Kiani, R. and M.N. Shadlen, *Representation of Confidence Associated with a Decision by Neurons in the Parietal Cortex.* Science, 2009. **324**(5928): p. 759-764.
24. Raposo, D., M.T. Kaufman, and A.K. Churchland, *A category-free neural population supports evolving demands during decision-making.* Nature Neuroscience, 2014. **17**(12): p. 1784-1792.
25. Najafi, F., et al., *Excitatory and Inhibitory Subnetworks Are Equally Selective during Decision-Making and Emerge Simultaneously during Learning.* Neuron, 2020. **105**(1): p. 165-+.
26. Licata, A.M., et al., *Posterior Parietal Cortex Guides Visual Decisions in Rats.* Journal of Neuroscience, 2017. **37**(19): p. 4954-4966.
27. Harvey, C.D., P. Coen, and D.W. Tank, *Choice-specific sequences in parietal cortex during a virtual-navigation decision task.* Nature, 2012. **484**(7392): p. 62-68.

28. Tseng, S.Y., et al., *Shared and specialized coding across posterior cortical areas for dynamic navigation decisions.* Neuron, 2022. **110**(15).
29. Zhong, L., et al., *Causal contributions of parietal cortex to perceptual decision-making during stimulus categorization.* Nature Neuroscience, 2019. **22**(6): p. 963-+.
30. Hwang, E.J., et al., *History-based action selection bias in posterior parietal cortex.* Nat Commun, 2017. **8**(1): p. 1242.
31. Hwang, E.J., et al., *Corticostriatal Flow of Action Selection Bias.* Neuron, 2019. **104**(6): p. 1126-+.
32. Morcos, A.S. and C.D. Harvey, *History-dependent variability in population dynamics during evidence accumulation in cortex.* Nature Neuroscience, 2016. **19**(12): p. 1672-1681.
33. Akrami, A., et al., *Posterior parietal cortex represents sensory history and mediates its effects on behaviour.* Nature, 2018. **554**(7692): p. 368-+.
34. Otazu, G.H., et al., *Engaging in an auditory task suppresses responses in auditory cortex.* Nat Neurosci, 2009. **12**(5): p. 646-54.
35. Bathellier, B., et al., *A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice.* Proceedings of the National Academy of Sciences, 2013. **110**(49): p. 19950-19955.
36. Lyamzin, D. and A. Benucci, *The mouse posterior parietal cortex: Anatomy and functions.* Neuroscience Research, 2019. **140**: p. 14-22.
37. Constantinople, C.M., et al., *Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases.* Elife, 2019. **8**.
38. Kobak, D., et al., *Demixed principal component analysis of neural population data.* eLife, 2016. **5**: p. e10989.
39. Worthy, D.A. and W. Todd Maddox, *A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W.K. Estes.* Journal of Mathematical Psychology, 2014. **59**: p. 41-49.
40. Jang, A.I., et al., *Positive reward prediction errors during decision-making strengthen memory encoding.* Nature Human Behaviour, 2019. **3**(7): p. 719-732.
41. Rouhani, N. and Y. Niv, *Signed and unsigned reward prediction errors dynamically enhance learning and memory.* eLife, 2021. **10**: p. e61077.
42. Sutton, R.S. and A.G. Barto, *Reinforcement learning: an introduction*. Second edition ed. Adaptive computation and machine learning series. 2018, Cambridge, Massachusetts: The MIT Press. 526.
43. Bathellier, B., et al., *A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(49): p. 19950-19955.
44. Fusi, S., et al., *A Neural Circuit Model of Flexible Sensorimotor Mapping: Learning and Forgetting on Multiple Timescales.* Neuron, 2007. **54**(2): p. 319-333.
45. Trepka, E., et al., *Entropy-based metrics for predicting choice behavior based on local response to reward.* Nature Communications, 2021. **12**(1): p. 6567.
46. Steiner, J. and C. Stewart, *Perceiving Prospects Properly.* American Economic Review, 2016. **106**(7): p. 1601–31.
47. Grossman, C.D., B.A. Bari, and J.Y. Cohen, *Serotonin neurons modulate learning rate through uncertainty.* Current Biology, 2022. **32**(3): p. 586-+.

48. Behrens, T.E.J., et al., *Learning the value of information in an uncertain world.* Nature Neuroscience, 2007. **10**(9): p. 1214-1221.
49. Yu, A.J. and P. Dayan, *Uncertainty, neuromodulation, and attention.* Neuron, 2005. **46**(4): p. 681-692.
50. Han, S. and F. Helmchen, *Behavior-relevant top-down cross-modal predictions in mouse neocortex.* Nature Neuroscience, 2024.
51. Driscoll, L.N., et al., *Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex.* Cell, 2017. **170**(5): p. 986-+.

Figure 1

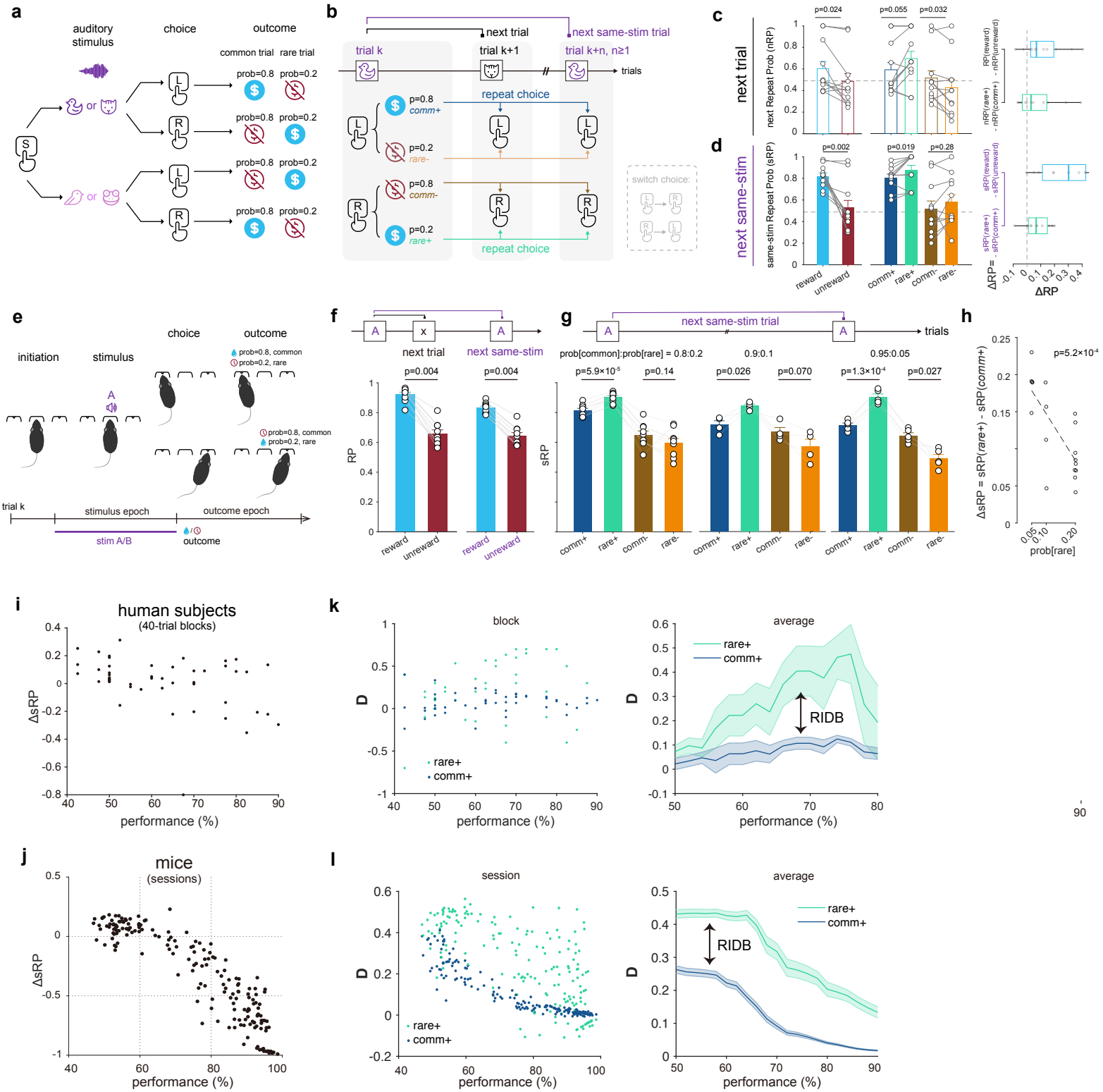

a
auditory stimulus
choice
outcome
common trial
rare trial
prob=0.8
prob=0.2
b
next trial
next same-stim trial
trial k
trial k+1
trial k+n, n≥1
trials
p=0.8 comm+
p=0.2 rare-
p=0.8 comm-
p=0.2 rare+
repeat choice
switch choice:
c
next trial
next Repeat Prob (nRP)
p=0.024
p=0.055
p=0.032
RP(reward) - nRP(unreward)
nRP(rare+) - nRP(comm+)
d
next same-stim
same-stim Repeat Prob (sRP)
p=0.002
p=0.019
p=0.28
sRP(reward) - sRP(unreward)
sRP(rare+) - sRP(comm+)
reward
unreward
comm+
rare+
comm-
rare-
ΔRP
e
initiation
stimulus
choice
outcome
prob=0.8, common
prob=0.2, rare
trial k
stimulus epoch
outcome epoch
stim A/B
outcome
f
next trial
next same-stim
p=0.004
p=0.004
RP
g
next same-stim trial
trials
prob[common]:prob[rare] = 0.8:0.2
0.9:0.1
0.95:0.05
p=5.9×10^-5
p=0.14
p=0.026
p=0.070
p=1.3×10^-4
p=0.027
sRP
h
p=5.2×10^-4
ΔsRP = sRP(rare+) - sRP(comm+)
prob[rare]
i
human subjects
(40-trial blocks)
ΔsRP
performance (%)
k
block
average
D
rare+
comm+
RIDB
performance (%)
j
mice
(sessions)
ΔsRP
performance (%)
l
session
average
D
rare+
comm+
RIDB
performance (%)

Figure 2

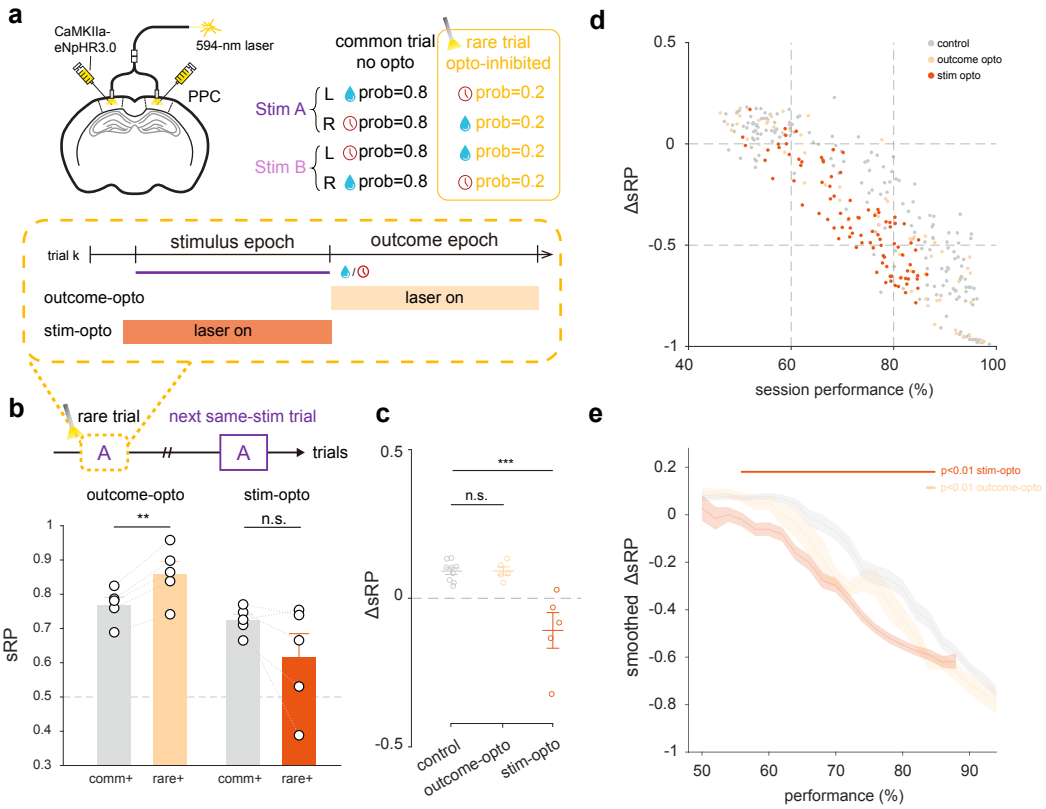

a
CaMKIIa-
eNpHR3.0
594-nm laser
PPC
common trial
no opto
rare trial
opto-inhibited
Stim A
L prob=0.8
R prob=0.8
prob=0.2
prob=0.2
Stim B
L prob=0.8
R prob=0.8
prob=0.2
prob=0.2
trial k
stimulus epoch
outcome epoch
outcome-opto
laser on
stim-opto
laser on
b
rare trial
next same-stim trial
A
A
trials
outcome-opto
stim-opto
**
n.s.
sRP
comm+
rare+
comm+
rare+
c
***
n.s.
ΔsRP
control
outcome-opto
stim-opto
d
control
outcome opto
stim opto
ΔsRP
session performance (%)
e
p<0.01 stim-opto
p<0.01 outcome-opto
smoothed ΔsRP
performance (%)

Figure 3

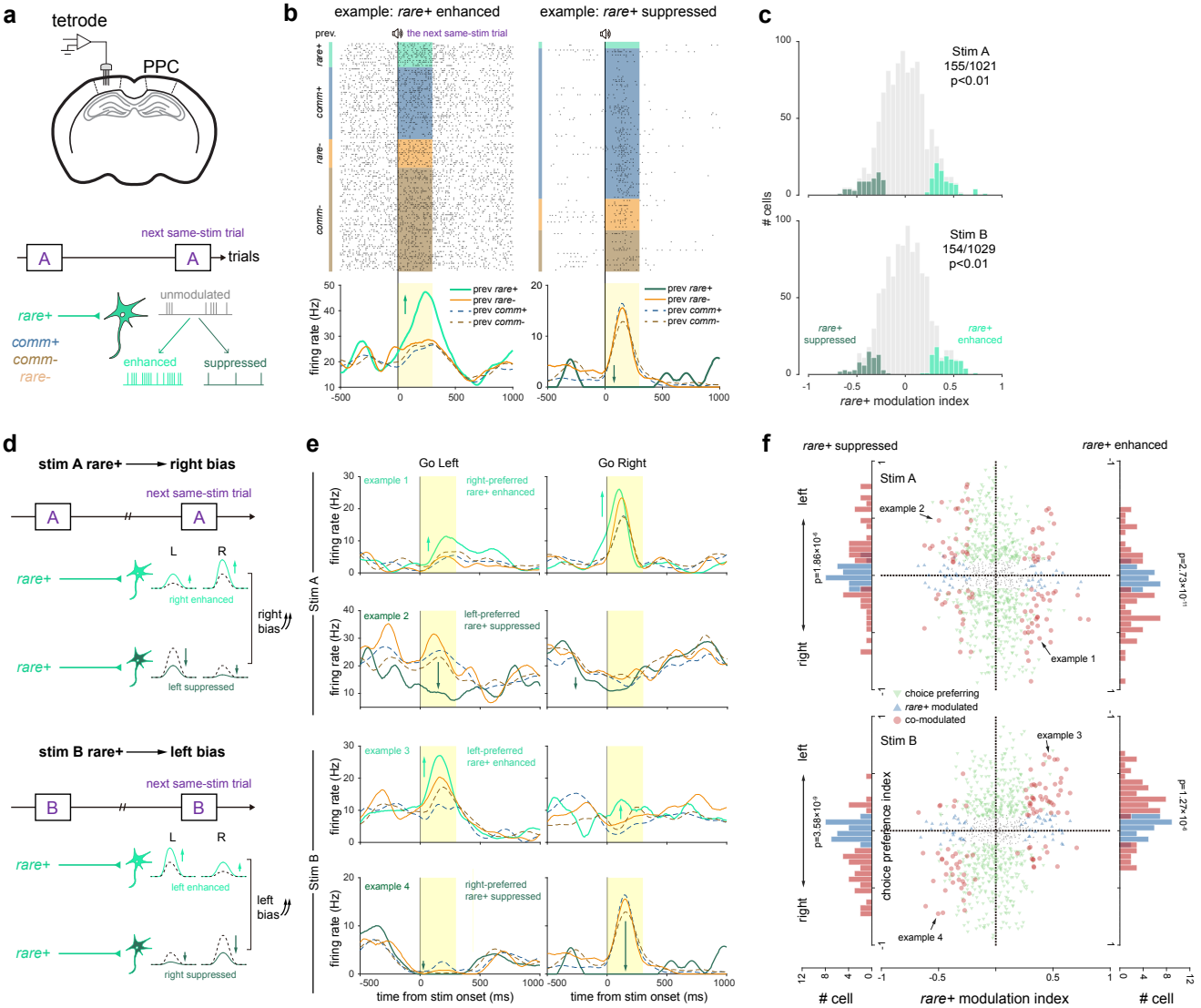

a
tetrode
PPC
next same-stim trial
A
A
trials
rare+
comm+
comm-
rare-
unmodulated
enhanced
suppressed
b
example: rare+ enhanced
example: rare+ suppressed
prev.
the next same-stim trial
rare+
comm+
rare-
comm-
firing rate (Hz)
prev rare+
prev rare-
prev comm+
prev comm-
c
Stim A
155/1021
p<0.01
Stim B
154/1029
p<0.01
cells
rare+ suppressed
rare+ enhanced
rare+ modulation index
d
stim A rare+ → right bias
next same-stim trial
A
A
L
R
rare+
right enhanced
right bias
rare+
left suppressed
stim B rare+ → left bias
next same-stim trial
B
B
L
R
rare+
left enhanced
left bias
rare+
right suppressed
e
Go Left
Go Right
Stim A
Stim B
example 1
right-preferred rare+ enhanced
example 2
left-preferred rare+ suppressed
example 3
left-preferred rare+ enhanced
example 4
right-preferred rare+ suppressed
firing rate (Hz)
time from stim onset (ms)
f
rare+ suppressed
rare+ enhanced
Stim A
Stim B
example 1
example 2
example 3
example 4
left
right
p=1.86×10⁻⁴
p=2.73×10⁻¹¹
p=3.58×10⁻⁹
p=1.27×10⁻⁶
choice preferring
rare+ modulated
co-modulated
choice preference index
cell
rare+ modulation index

Figure 4

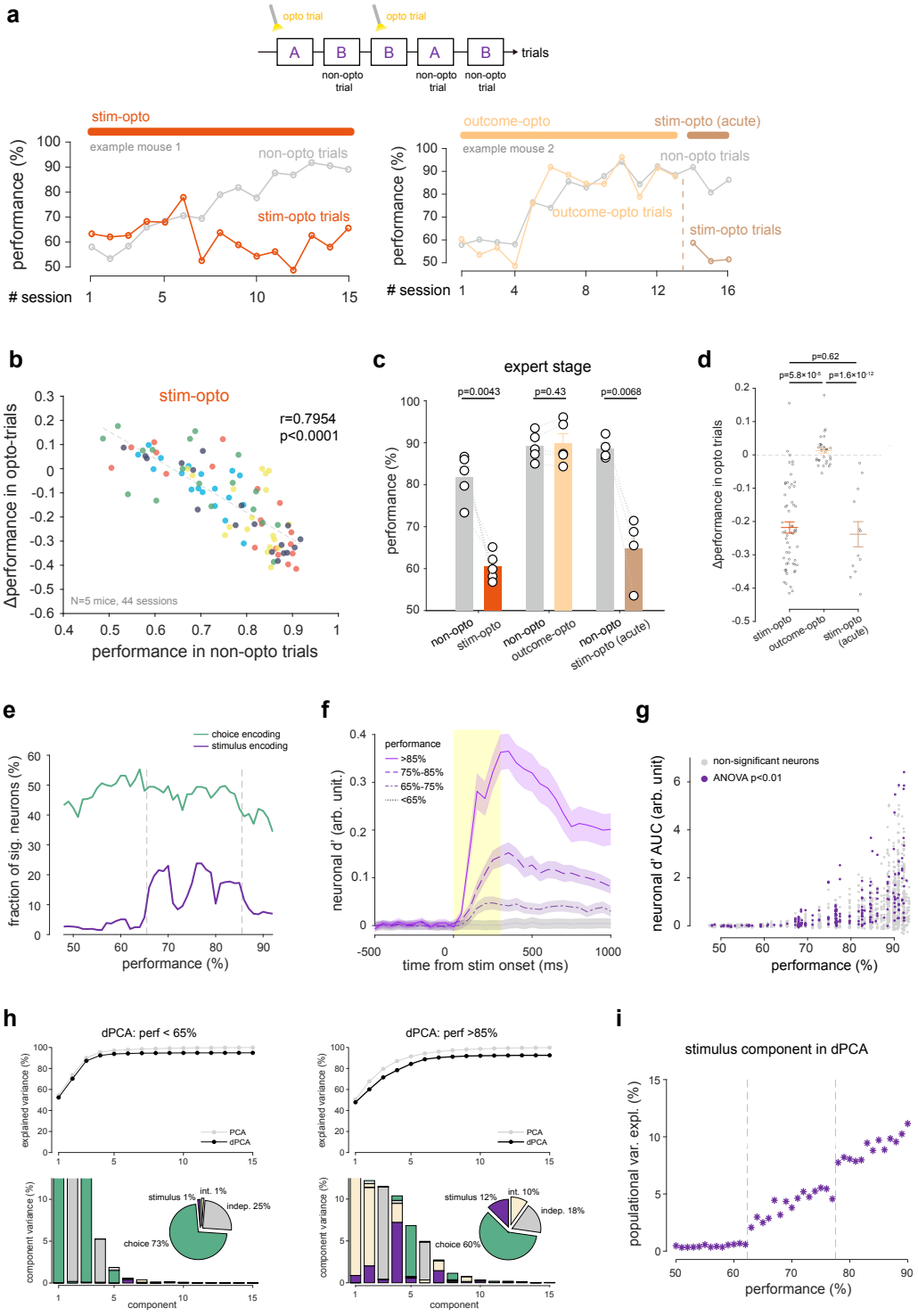
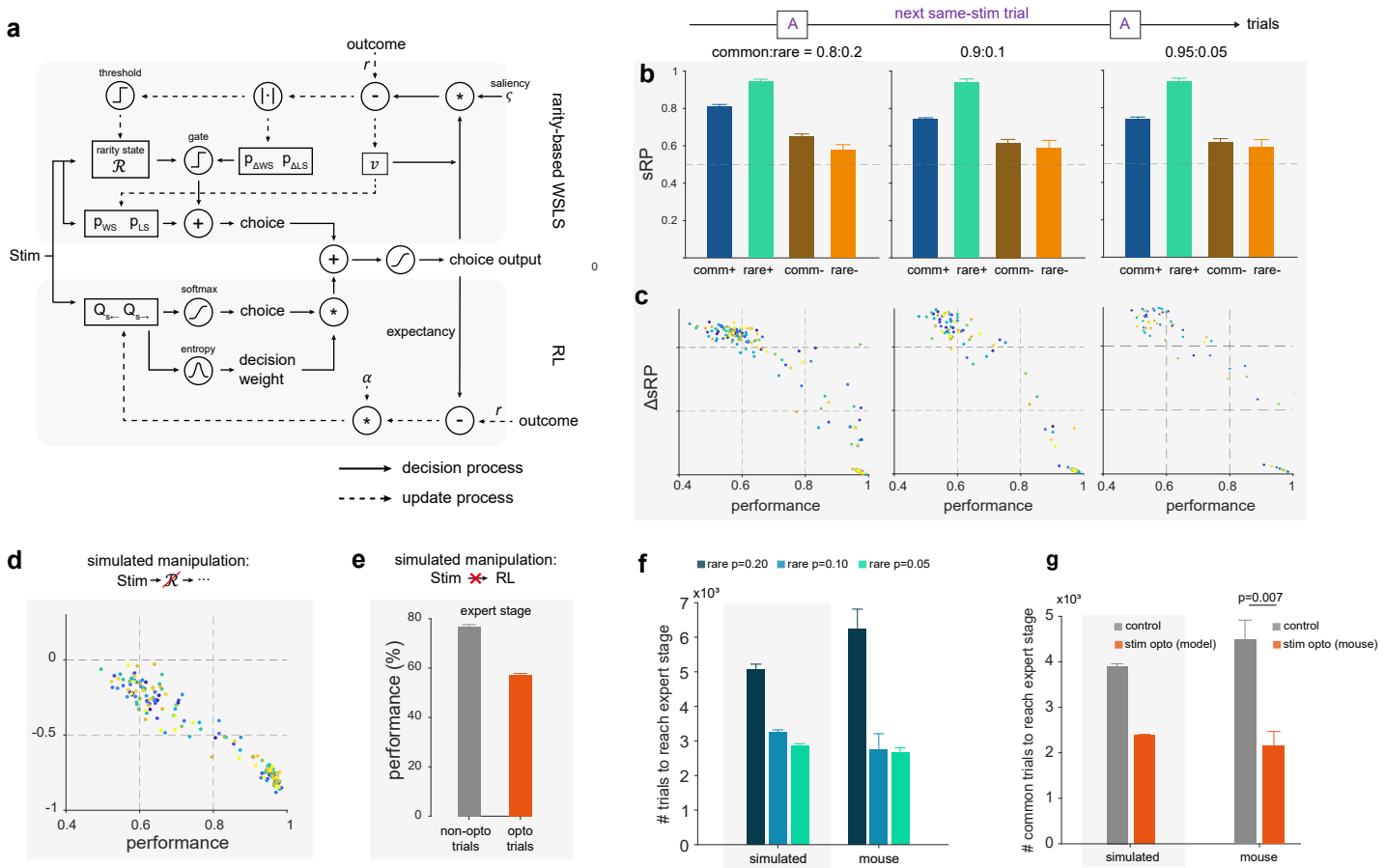

a
opto trial
opto trial
A
B
B
A
B
trials
non-opto trial
non-opto trial
non-opto trial
stim-opto
example mouse 1
non-opto trials
stim-opto trials
performance (%)
session
outcome-opto
stim-opto (acute)
example mouse 2
non-opto trials
outcome-opto trials
stim-opto trials
b
stim-opto
r=0.7954
p<0.0001
N=5 mice, 44 sessions
Δperformance in opto-trials
performance in non-opto trials
c
expert stage
p=0.0043
p=0.43
p=0.0068
performance (%)
non-opto
stim-opto
non-opto
outcome-opto
non-opto
stim-opto (acute)
d
p=0.62
p=5.8×10⁻⁵
p=1.6×10⁻¹²
Δperformance in opto trials
stim-opto
outcome-opto
stim-opto (acute)
e
choice encoding
stimulus encoding
fraction of sig. neurons (%)
performance (%)
f
performance
>85%
75%-85%
65%-75%
<65%
neuronal d' (arb. unit.)
time from stim onset (ms)
g
non-significant neurons
ANOVA p<0.01
neuronal d' AUC (arb. unit)
performance (%)
h
dPCA: perf < 65%
explained variance (%)
PCA
dPCA
component variance (%)
stimulus 1%
int. 1%
indep. 25%
choice 73%
component
dPCA: perf >85%
stimulus 12%
int. 10%
indep. 18%
choice 60%
i
stimulus component in dPCA
populational var. expl. (%)
performance (%)

Figure 5


a
outcome
threshold
saliency
rarity state
gate
choice
Stim
choice output
softmax
choice
expectancy
entropy
decision weight
outcome
rarity-based WSLS
RL
decision process
update process
next same-stim trial
trials
common:rare = 0.8:0.2
0.9:0.1
0.95:0.05
b
sRP
comm+ rare+ comm- rare-
c
ΔsRP
performance
d
simulated manipulation:
Stim → ℛ → ···
performance
e
simulated manipulation:
Stim ✗ RL
expert stage
performance (%)
non-opto trials
opto trials
f
rare p=0.20
rare p=0.10
rare p=0.05
trials to reach expert stage
simulated
mouse
g
p=0.007
control
stim opto (model)
stim opto (mouse)
common trials to reach expert stage
simulated
mouse