Separable neurocomputational mechanisms underlying multisensory learning

Saurabh Bedi*[1,2], Ella Casimiro*[1,2], Gilles de Hollander[1,2], Nina Raduner[2,4,6], Fritjof Helmchen[2,4,5], Silvia Brem[2,4,6], Arkady Konovalov[3], and Christian C. Ruff[1,2,4]

[1]Zurich Center for Neuroeconomics (ZNE), University of Zurich
[2]University Research Priority Program (URPP), Adaptive Brain Circuits in Development and Learning (AdaBD), University of Zurich
[3]Centre for Human Brain Health, School of Psychology, University of Birmingham
[4]Neuroscience Center Zurich, ZNZ
[5]Brain Research Institute, University of Zurich
[6]Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital of Psychiatry Zurich, University of Zurich

November 17, 2025

* Saurabh Bedi and Ella Casimiro are co-first authors

Abstract

Efficient control of behavior requires multisensory learning from information distributed across senses. However, most neurocomputational studies have focused on unisensory signals. Here, we identify distinct but interacting neurocomputational mechanisms that support learning of multisensory associations. We designed a task in which behaviorally relevant information was available only from combinations of visual cues with either auditory or tactile cues. In 58 participants undergoing fMRI, we dissociated three processes: multisensory statistical learning (SL), modeled as stimulus-locked Shannon surprise; reinforcement learning (RL), modeled as feedback-locked signed reward prediction errors (RPEs); and feedback-locked unsigned RPEs (uRPEs), reflecting surprise about reward outcomes. Behaviorally, response times scaled with Shannon surprise (SL) while accuracy improved with feedback (RL). Model-based fMRI revealed dissociable but complementary networks: RPEs engaged ventral striatum, vmPFC, and left angular gyrus; surprise recruited bilateral angular gyrus, dlPFC, and precuneus; and uRPEs involved insula, dorsomedial prefrontal, and lateral frontoparietal cortices. Several of these regions are not typically implicated in unisensory studies, suggesting specialization for multisensory learning. All three networks were modality-general, i.e., they showed comparable strength for audiovisual and visuotactile learning. Notably, left angular gyrus tracked both Shannon surprise and RPE, identifying it as a potential hub for integrating structural and value information. These findings reveal that the brain engages distinct but complementary systems for structure-based, reward-based, and outcome-surprise computations. By combining behavioral modeling and fMRI with a novel task design, we provide a principled framework for dissecting the neurocomputational architecture of multisensory learning.

Introduction

Most events in daily life engage multiple senses simultaneously: Visual stimuli are paired with sounds, scents, or tactile sensations that the brain seamlessly integrates into a coherent multisensory experience. Reflecting this, neuroscience has adopted the idea of a multisensory brain [1, 2, 3, 4, 5], supported by evidence that even brain regions traditionally viewed as unisensory respond to inputs from different sensory modalities [6]. Multisensory illusions, such as the ventriloquist and the McGurk effect, demonstrate that this integration is automatic and can fundamentally alter perception [7]. While much is known about how the brain integrates multisensory signals in perception [8, 9], far less is understood about how it learns multisensory associations, particularly when

stimulus and action-relevant information resides exclusively in the combination of sensory modalities rather than in any individual sense. Understanding the computational and neural principles of multisensory learning has so far remained a gap in our general knowledge about learning.

This gap has broad implications. Multisensory learning underlies the foundations of many real-world behaviors and cognitive development [10, 11]. For instance, the development of flavor preferences in childhood shapes lifelong eating habits [12] and relies on the integration of smell, taste, and touch, as no single modality suffices to create the flavor experience [13, 14]. Language acquisition similarly depends on multisensory learning [11]: reading involves associating visual letters with auditory phonemes and writing involves coordination between visual and tactile information. Impairments in these multisensory processes have been associated with developmental dyslexia [15, 16] , a learning disability affecting roughly 5 to 10 percent of the population [17]. These examples underscore the deep reliance of core cognitive functions on multisensory learning, highlighting the need to investigate its underlying neural and computational mechanisms.

Understanding the computational mechanisms underlying multisensory learning begins with identifying the core goals that guide learning more generally. Broadly, learning is shaped by two distinct computational objectives that are each tied to a different source of information. **Reinforcement learning** (**RL**) involves learning to predict rewarding stimuli (Pavlovian learning) [18] or rewarding actions (Instrumental learning) [19] with the goal of maximizing cumulative reward [20]. In contrast, **statistical learning** (**SL**) refers to the learning of environmental regularities, with the goal of enabling efficient allocation of limited cognitive resources to frequently occurring stimuli, as formalized for example by the efficient coding hypothesis [21, 22, 23]. These computations are conceptually distinct: RL is driven by rewards, while SL is driven by the statistical structure of the environment. Here, we examine how each of these learning computations operate in multisensory learning, where both structure and reward are embedded across sensory modalities rather than within unisensory information.

Disentangling RL and SL in multisensory contexts poses a fundamental challenge because these processes often interact. While SL can operate independently of reinforcement, as in classic latent learning [24], it can also support RL by uncovering structure not only in sensory states but also in transitions, actions, and policies, thus reducing the computational complexity of model-based reinforcement learning through compressed representations [25, 26, 27, 28, 29]. Here, we investigate both computations simultaneously in the context of multisensory learning, at both the behavioral level (using computational modeling) and the neural level (using functional magnetic resonance imaging). We minimize the issue of the possible interactions between the two computations by designing a task that keeps RL and SL functionally distinct, enabling a principled investigation of their individual contributions. This provides a foundation for future work on how these computations interact when structure and reward co-occur.

We designed a novel task to orthogonalize SL and RL in multisensory contexts by separating their informational sources and temporal dynamics. Participants viewed nine unique multisensory pairings (3 images $\times$ 3 sounds or 3 tactile patterns) and judged whether the pairing was correct or not. To facilitate comprehension, the task was embedded in a cover story: Participants took the role of an animal scientist and viewed images of insects while hearing possible calls or feeling tactile patterns of particular dances that may or may not attract a mate for this species. Correct responses ("attracts" or "does not attract") depended only on the combination of cues, never on a single modality. Probabilistic feedback (+1 for correct and 0 for incorrect judgments) required participants to rely on RL (learning through trial-and-error) to learn which pairings predicted success in attracting a mate. In parallel, SL was embedded orthogonally in the frequency of presenting specific multisensory combinations, providing statistical structure that was irrelevant for reward (see Methods for details). We modeled SL as reward-independent **Shannon surprise** [30] at stimulus onset [31, 32, 33, 34]. Independently, we modeled RL using **reward prediction errors** (**RPEs**), learning signals computed at feedback that capture the discrepancy between expected value estimates and actual rewards. For each participant, RPEs were derived from the best-fitting model among several RL models reflecting plausible task strategies. Model comparison results are reported in the Methods (showing that three of five candidate models provided good fits, with the majority of participants best described by a basic Q-learning model). Along with the statistical regularities in the stimuli, participants' learning also created an emergent statistical regularity in reward outcomes. To capture this, we included **unsigned RPEs (uRPEs)** which are feedback-driven signals that index outcome surprise irrespective of valence. uRPE thus sits at the intersection of RL and SL: like RPE, it depends on feedback, but like SL, it quantifies statistical surprise. Together, Shannon surprise, signed RPE, and uRPE provided complementary measures to dissect the distinct parallel computations underlying multisensory learning.

This computational and temporal dissociation allowed us to investigate the neural signatures

of SL and RL in multisensory learning. Because SL-related surprise occurs at stimulus onset while RL-related RPEs and uRPEs occur at feedback, we could temporally separate structure-based from reward-based learning signals in trial-wise fMRI analyses. Previous research in purely unisensory contexts has associated statistical surprise with activity in frontal regions (e.g., anterior insula, inferior, medial, and superior frontal gyri), posterior parietal cortex (including the superior and inferior parietal lobules), striatal regions, temporal cortices, and anterior and posterior cingulate cortex [34, 35, 36, 37]. Similarly, in unisensory contexts, signed RPEs are typically linked to the ventral striatum, medial frontal cortex, hippocampus, anterior insula, cingulate cortex, and superior frontal regions [37, 38, 39], whereas uRPEs have been associated with activity in the striatum, insula, cingulate cortex, dorsolateral prefrontal cortex, inferior parietal lobule, and temporal regions [39, 37]. However, because these findings are largely derived from unisensory paradigms, it remains unknown whether, where, and how the corresponding computations are implemented when learning depends on associations distributed across multiple senses.

From a purely theoretical perspective, accounts of multisensory processing propose several possible architectures for how multisensory computations could be implemented in the brain. One influential framework suggests that specific cortical regions may act as multisensory hubs that integrate information across sensory modalities [6]. These hubs may be modality-specific - i.e., engaged differently depending on which specific multisensory combinations are processed - or modality-general, operating similarly across various pairings of different senses. To address this question, we tested these possibilities directly in the context of multisensory learning, by examining Shannon surprise, signed RPEs, and uRPEs across two distinct multisensory contexts: audiovisual and visuotactile. Specifically, we asked three key questions: (1) which brain regions are selectively engaged in multisensory learning beyond canonical unisensory learning systems, (2) are these learning signals modality-specific or modality-general across sensory contexts, and (3) are the different computational components of learning (structure-based (SL), reward-based (RPE), and outcome-surprise–based (uRPE)) supported by shared or distinct neural mechanisms. By combining trial-wise computational modeling with temporally separated stimulus and feedback phases, our design allowed us to isolate each computation and directly test these hypotheses within a unified experimental framework.

Results

The experiment

To separate SL and RL during multisensory learning, we developed a novel two-alternative forced choice task performed by participants (n = 64) across six blocks in an fMRI scanner. In each block, they were presented with one of two multisensory stimulus sets (either audiovisual or visuotactile), each comprising nine unique pairings of three insect images with either three auditory "calls" (beep sequences) or three tactile "dances" (vibration patterns). The participants' goal was to infer whether each multisensory pairing would attract a mate, mimicking a scientist classifying courtship success (Fig. 1a, b). Critically, the correct response depended only on the specific multisensory combination, not on any individual unisensory element (Fig. 1c). Feedback was probabilistic: correct pairings led to successful attraction (+1 reward) on most trials, but occasionally produced failed attraction (0 reward) (see Methods). Conversely, incorrect pairings usually led to failed attraction (0 reward), but sometimes resulted in successful attraction (+1 reward). This probabilistic feedback ensured that the task remained challenging, promoting continuous RPE signals due to reinforcement learning throughout each block. Participants were compensated with real money based on cumulative rewards they earned during the task.

To isolate SL, we manipulated the frequency with which each multisensory stimulus appeared, embedding statistical structure only in the combinations, not in the visual, auditory or tactile components individually (Fig. 1d). SL was not required to perform the task or obtain rewards. In contrast, RL was essential: participants had to learn and indicate which combinations led to successful mating. Critically, both successful (green squares) and non-successful (white squares) combinations occurred equally often (50%), meaning that each response (success vs. failure) was correct half the time across all stimuli. This ensured that SL provided no informational advantage for RL. We additionally modeled a third computational signal: the uRPE, which captures the magnitude of outcome surprise regardless of reward valence. While uRPE is computed at feedback alongside signed RPE, it reflects a different computational quantity: the unpredictability of the reward itself. In this sense, uRPE lies at the intersection of RL and SL: like RPE, it is grounded

in feedback and value learning, yet like Shannon surprise, it captures statistical unexpectedness, although here, in the distribution of reward outcomes. By dissociating these computations in both information content and temporal structure, with SL relevant surprise occurring at stimulus onset and RL relevant RPE/uRPE arising after choice (Fig. 1b), our design enabled independent modeling and neural analysis of SL and RL of truly multisensory stimulus and stimulus-response associations.

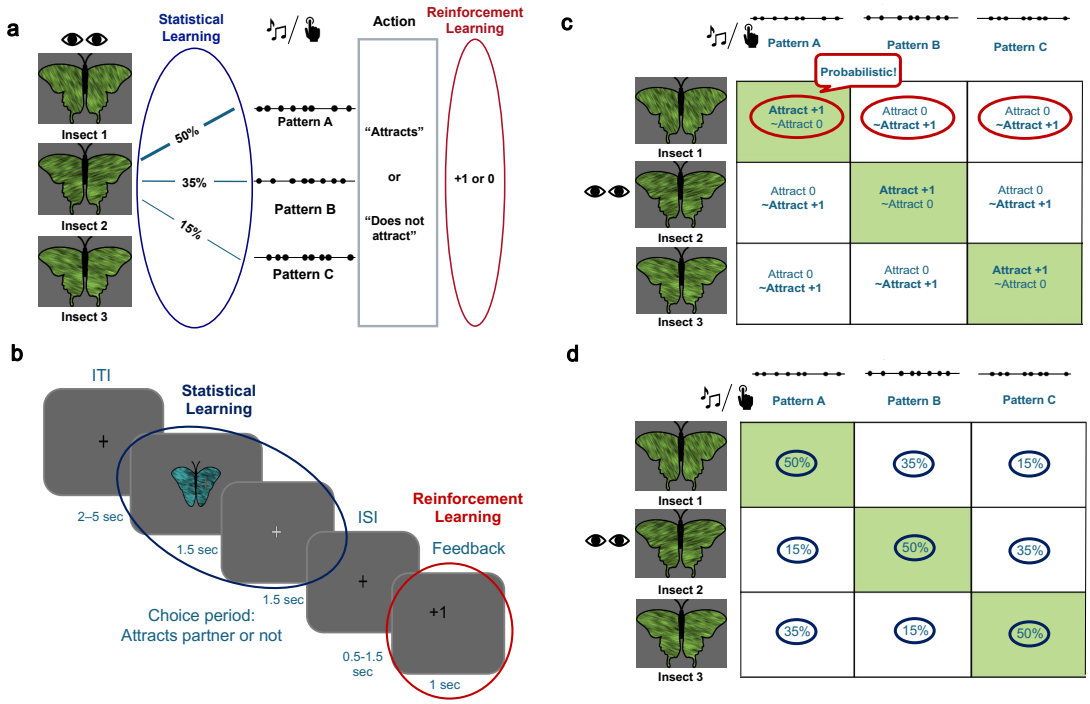

Figure 1: **Task design**. a, Schematic of the multisensory learning task. Participants acted as "insect scientists" learning which multisensory pairings (image × sound/touch) predicted mating success. b, Trial timing showing separate epochs for stimulus presentation (driving statistical learning) and feedback (driving reinforcement learning). c, Reinforcement structure: probabilistic outcomes (green = "attracts" is rewarded, white = "not attracts" is rewarded) ensured ongoing prediction errors. d, Statistical structure: frequency distribution of multisensory pairings embedding regularities only at the combination level. Together, the task orthogonalized statistical and reinforcement learning signals in both temporal and informational domains.

Models of learning

Reinforcement Learning

We modeled RL to quantify how participants used probabilistic feedback to adapt their (Attract/Don't attract) choices for each multisensory pairing, and to derive trial-by-trial learning signals (RPE and uRPE) for neural analyses. Because our task was designed to comprise only fully instructed transitions between known stimuli and states, participants did not need to learn a model of the testing environment. This minimized potential overlap between the statistical structure and reward contingencies. We used the **Q-learning algorithm** as our base model, in which action values (Q-values) are updated based on the discrepancy between expected value estimates and actual rewards, which is known as the reward prediction error:

$$Q_c(t) = Q_c(t-1) + \alpha * (R - Q_c(t-1))$$

where $Q_c(t)$ is the Q-value for the chosen action at trial t, R is the received reward (1 or 0), and α is the learning rate that determines how strongly RPEs update Q-values. In addition to computing signed RPEs, we also derived the uRPE as the absolute value of the RPE. In our task, uRPE reflects task-relevant unpredictability in feedback, providing a distinct, non-valenced measure of reward-based surprise for neural analysis. It has been shown in purely unisensory

contexts that RPEs and uRPEs engage partially distinct neural substrates: while both involve midbrain, striatal, and insular regions, unsigned PEs more consistently activate the cerebellum, dorsolateral and dorsomedial prefrontal cortex, supplementary motor area, supramarginal gyrus, and posterior parietal and temporal cortices, regions implicated in attentional control and salience processing [39, 37]. Including both signals in our analysis therefore allows us to dissociate neural systems involved in value updating from those involved in detecting surprising reward outcomes during multisensory learning.

The **basic Q-learning** model assumes a constant learning rate across all trials [20, 40]. However, participants may have various biases or adopt certain strategies depending on prior expectations or task instructions. To capture these individual differences, we implemented several model variants of Q-learning inspired by the prior literature:

1. Asymmetric Learning Rates (**Asym**): Evidence suggests that participants may learn differently from positive and negative RPEs [41, 42, 43, 44]. To capture this asymmetric learning, we implemented models with distinct learning rates for positive and negative RPEs.

2. Transfer Learning Models (**Transfer**): The task instructions explicitly stated that each insect had only one correct call or dance. This means that if a participant received a reward for pairing insect A with call 1 (A1), they could infer that other pairings involving the same insect (e.g., A2, A3) or the same call (e.g., B1, C1) are likely incorrect. This rule-based structure, which was available from the instructions but not through direct feedback, creates opportunities for inference beyond the presented trial. We modeled this as a structured "transfer" mechanism that allowed RPEs to update not only the chosen option but also unchosen but logically related options [45].

3. Initial Bias Model (**Vinit**): Participants might begin the task with preexisting biases toward one response option (e.g., left/right). We modeled this by assigning different initial values to actions, allowing for baseline preferences that could influence early behavior.

We also fit combinations of these models (e.g., Transfer+V_{init}), yielding eight candidate models in total (see Methods). Of these, five showed acceptable parameter recovery ($r > 0.6$) and were used for final model selection. Across participants, 31 were best fit by the basic Q-learning model, 16 by the Transfer model, and 11 by the Asymmetric model, while neither the Vinit model nor the Transfer+V_{init} combination best fit any participant (see Methods for details). Each participant's best-fitting model per block was then used to compute trial-by-trial estimates of RPE and uRPE for neural analyses.

Action selection on each trial was modeled using a softmax function:

$$P_c(t) = \frac{e^{\beta * Q_c(t)}}{e^{\beta * Q_c(t)} + e^{\beta * Q_u(t)}}$$

where Q_c is the value of the chosen response and Q_u is the value of the unchosen response. The inverse temperature parameter β scales the level of randomness in choices and was fit for each individual for each block.

Statistical Learning

We modeled SL to test whether participants implicitly tracked the frequency structure of multisensory combinations, despite its irrelevance for reward, and to obtain a stimulus-locked, trial-by-trial index of structure-driven surprise for behavioral and fMRI analyses. We used a simple Bayesian observer that maintains beliefs over the nine combinations (uniform prior at block start) and updates them from observed frequencies [31, 32, 33, 34]:

$$p_i(t) = \frac{1 + \sum_{t'=1}^{t} I_{[o(t')=i]}}{9 + t},$$

where $p_i(t)$ is the belief that stimulus i occurs on trial t and $I_{[o(t')=i]} = 1$ if i was observed on trial t'. The SL regressor is the Shannon surprise at stimulus onset,

$$\text{Surprise}_o(t) = -\log\big(p_o(t)\big),$$

which quantifies how unexpected the presented combination *o(t)* is given current beliefs. Prior work has shown that such signals can modulate behavioral response times and neural responses, even in the absence of explicit rewards [34]. Because regularities were confined to joint combinations and were balanced across actions and modalities, this stimulus-locked signal is orthogonal to RL. In contrast to uRPE, which captures surprise about reward outcomes at feedback, Shannon surprise quantifies surprise about stimulus states at onset.

Behavioral results

Our behavioral data indicate that participants were engaged in both reinforcement and statistical learning during the task. As shown in Fig. 2a, accuracy increased steadily over trials in both audiovisual and visuotactile blocks, reflecting reinforcement learning from the reward feedback. The probabilistic reward structure encouraged continued exploration and prevented early learning saturation. To test the relationship between surprise and response times, we fit a mixed-effects regression model that included surprise, Q-values, and subject-level random intercepts as predictors. This analysis confirmed the hypothesized significant positive relationship between surprise and response time in both modalities (audiovisual: $\beta = 0.023$, one-sided p = 0.013; visuotactile: $\beta = 0.017$, one-sided p = 0.037). This effect, visualized in Fig. 2b, suggests that participants were sensitive to the statistical regularities of the environment, exhibiting slower responses to rare or unexpected stimuli which is a hallmark of implicit SL.

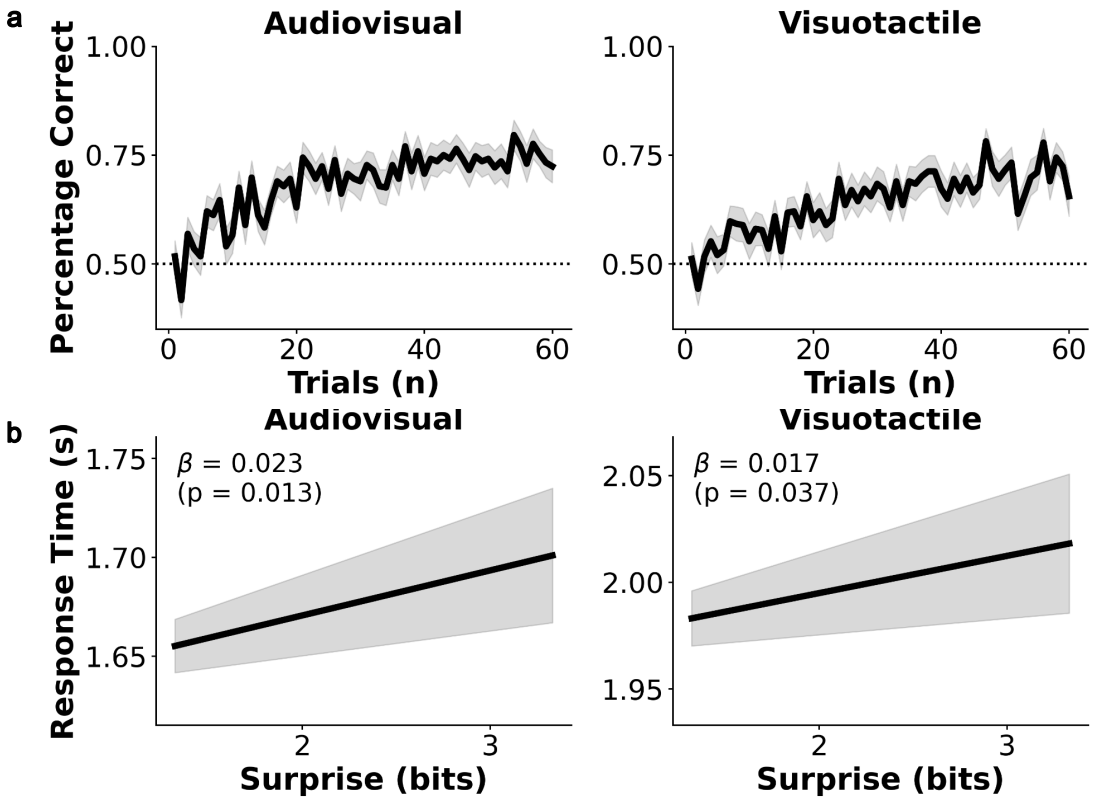

Figure 2: **Behavioral signatures of reinforcement and statistical learning.** a, Group-averaged learning curves showing accuracy across trials (smoothed, ±SEM) for audiovisual and visuotactile conditions. Increasing accuracy indicates reinforcement learning driven by feedback-based reward prediction errors. b, Response times as a function of Shannon surprise (computed at stimulus onset) showing the hypothesized longer latencies for rare combinations, consistent with implicit statistical learning. Each line represents the fixed-effect slope from a mixed-effects model controlling for subject-level random intercepts. Together, these effects confirm that participants engaged both feedback-based and structure-based learning mechanisms.

Neural underpinnings of the two types of multisensory learning

Having established distinct behavioral and computational signatures of RL and SL, we next examined whether these two learning processes are implemented in separate neural systems. Specifically, we tested three key questions: (1) whether SL and RL signals are encoded in distinct brain regions, (2) whether their neural substrates are modality-general or modality-specific, and (3) whether any of the identified brain areas appear to be selectively engaged in multisensory learning, in being distinct from those reported in the prior unisensory learning literature. To address these questions, we conducted a whole-brain model-based fMRI analysis [46], using trial-wise RPEs, uRPEs and Shannon surprise as parametric modulators for RL and SL. Standard nuisance regressors (motion, physiological noise) were also included in the general linear model, and all results were thresholded

using non-parametric cluster-based inference (SnPM13).

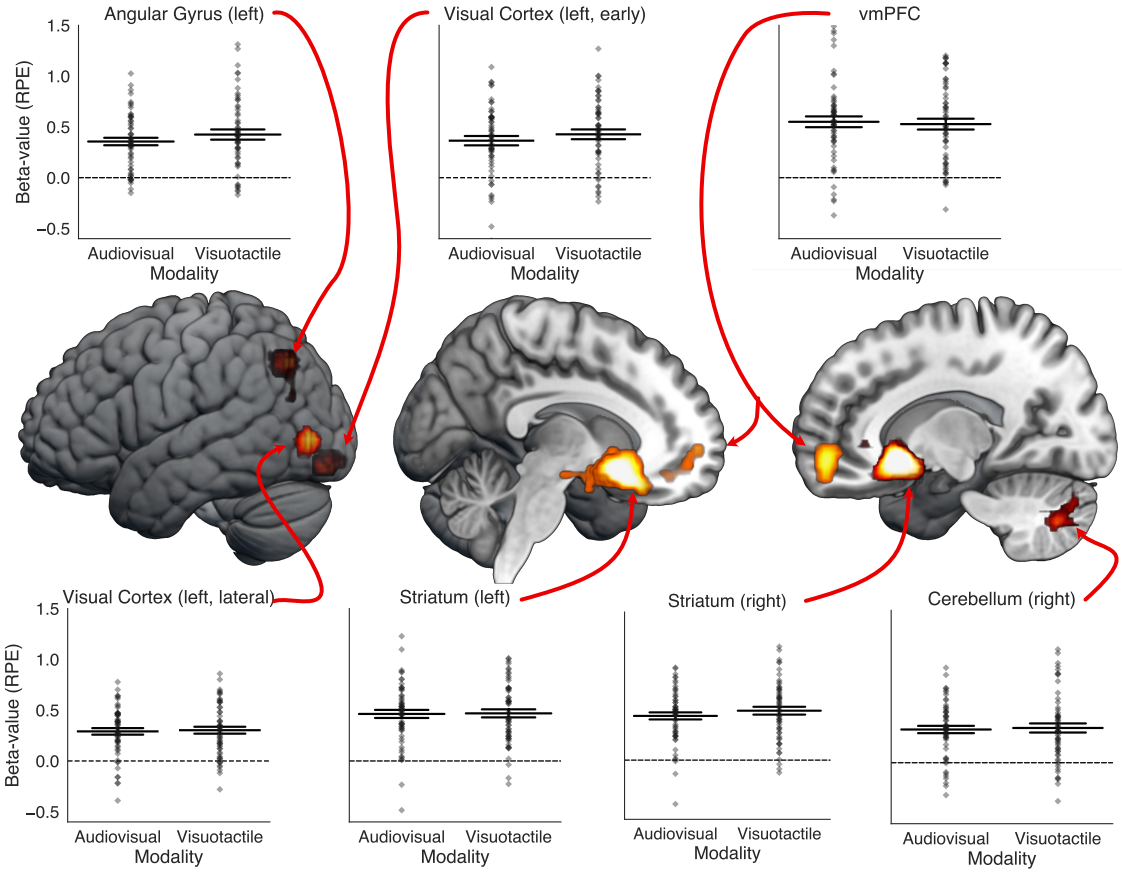

Figure 3: **Neural correlates of reward prediction errors (RPEs)**. The brain images in the center display regions where signed RPEs, modeled as a parametric modulator at feedback onset, positively correlated with BOLD activation. RPE reflects the signed difference between expected and received feedback (reward-dependent) indexing reinforcement learning. Statistical maps were thresholded at a cluster-forming $t > 8.0$ and inference was performed at cluster-level $p < .05$, FWE-corrected using SnPM, and rendered on the MNI152 standard brain. Significant activations were observed in the bilateral striatum, ventromedial prefrontal cortex (vmPFC), early lateral left visual cortex, left angular gyrus, and right cerebellum. The accompanying strip plots display the mean beta estimates of the RPE contrast across all voxels within the thresholded cluster in MNI space, with each diamond marker representing data from an individual participant. Estimates are shown separately for audiovisual and visuotactile learning, revealing that RPE-related responses did not differ significantly between modalities. The widest horizontal line indicates the group mean, whereas the more narrow lines indicate a distance of one standard error of the mean (SEM).

In our first whole-brain GLM analysis, we collapsed across audiovisual and visuotactile conditions to target domain-general learning signals (domain-specific signals are examined in a second set of analyses). These first analyses revealed that the three learning signals (RPE, Shannon surprise, and uRPE) were each associated with largely non-overlapping brain networks, supporting the idea that distinct neural architectures support reinforcement and statistical learning. Trial-by-trial RPEs during the feedback phase were correlated with activity in the ventromedial prefrontal cortex (vmPFC) and ventral striatum (caudate nucleus, putamen), all regions consistently implicated in unisensory reinforcement learning, particularly in value representation and reward prediction error encoding across modalities [47, 48, 49, 50, 51, 52]. Additional RPE-related activation appeared in visual association areas (fusiform, lingual, and middle occipital gyri; see Fig. 3 and Table 1), likely reflecting visual stimulus-dependent value modulation or attentional processing [53], as suggested by previous research that has consistently linked visual areas to RPEs [53, 37, 38]. Notably, activity in left angular gyrus also tracked RPEs, hinting at a potential region beyond classical reward circuitry. This region has so far only been reported to be involved in static reward anticipation [54, 55], but not in representation of dynamically evolving RPEs [37, 53, 38] . Inspection of the

reported coordinates and available activation maps from these meta-analyses revealed no spatial overlap with the present activation cluster, suggesting that its involvement in RPE processing here may reflect a specific engagement in multisensory learning.

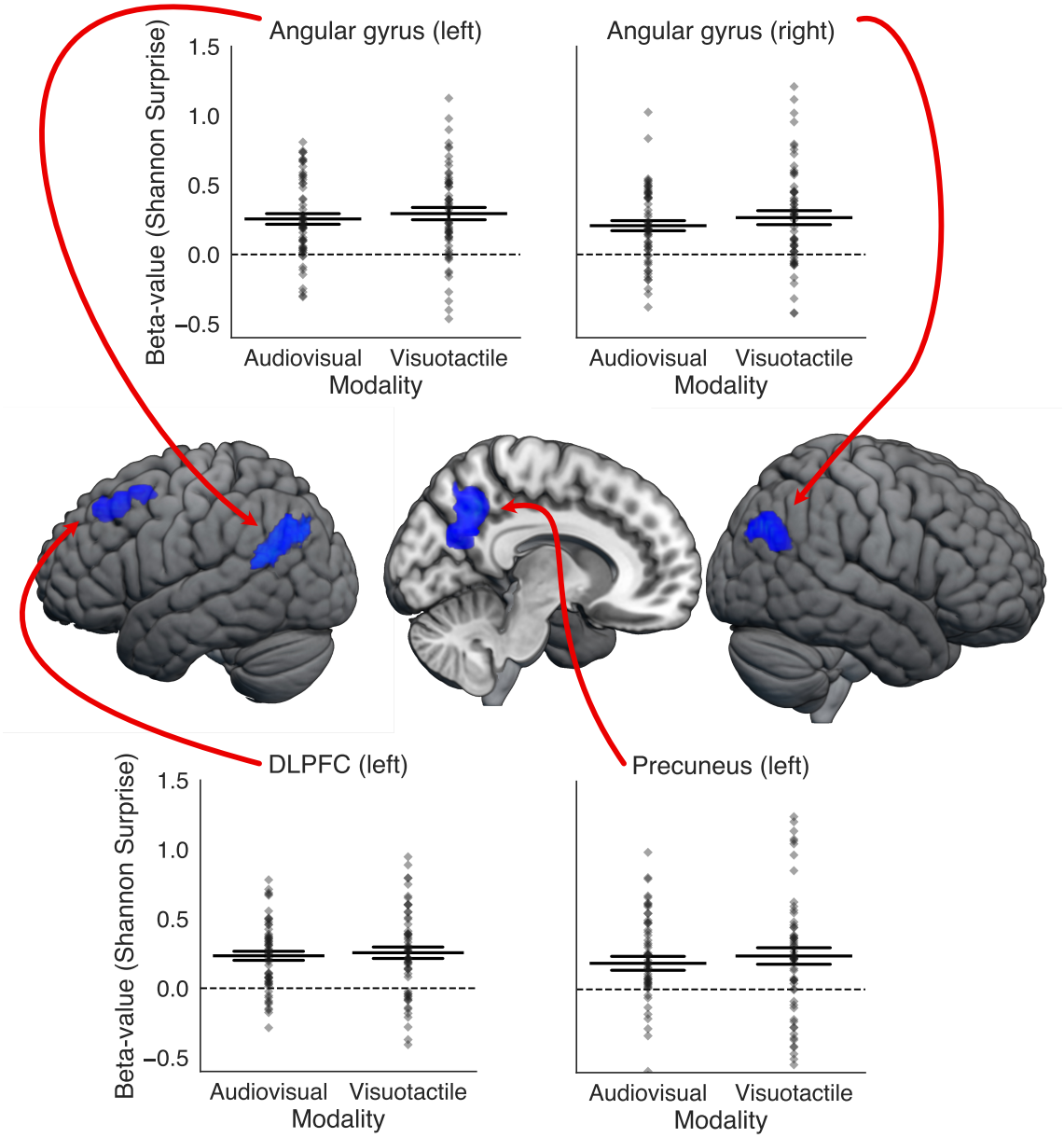

Figure 4: **Neural correlates of surprise**. The brain images in the center display regions where the Shannon surprise, modeled as a parametric modulator at stimulus onset, positively correlated with BOLD activation. The Shannon surprise quantifies the unexpectedness of the stimulus and indexes statistical learning. Statistical maps were thresholded at a cluster-forming $t > 3.1$ and inference was performed at $p < .05$, FWE-corrected using SnPM, and rendered on the MNI152 standard brain. Surprise significantly correlated with brain activation in the bilateral angular gyrus, left dorso-lateral prefrontal cortex (dlPFC), and left precuneus. The accompanying strip plots display the mean beta estimates of the RPE contrast across all voxels within the thresholded cluster in MNI space, with each diamond marker representing data from an individual participant. Estimates are shown separately for audiovisual and visuotactile learning, revealing that RPE-related responses did not differ significantly between modalities. The widest horizontal line indicates the group mean, whereas the more narrow lines indicate a distance of one standard error of the mean (SEM).

In line with the assumption that separate neural systems mediate RL vs SL, Shannon surprise during stimulus presentation engaged a separate set of regions, including bilateral angular gyri, the left precuneus, and the left dorsolateral prefrontal cortex (dlPFC; see Fig. 4 and Table 2). To

our knowledge, only a handful of studies have examined the neural correlates of Shannon surprise in unisensory contexts [37, 36, 35, 34] and no large-scale meta-analyses are available. Despite this limited evidence base, our findings partially converge with previous reports: all four previous studies also identified a parietal cluster near to (or extending into) the angular gyrus, and two of the four studies also linked the dlPFC to the Shannon surprise, albeit in the right hemisphere [37, 34]. Notably, none of these studies reported activation in the precuneus area identified here, suggesting that this region may provide a novel, modality-independent representation of multisensory surprise. In any case, the absence of spatial overlap with RPE-related regions further supports the conclusion that statistical and reinforcement learning rely on distinct neurocomputational pathways.

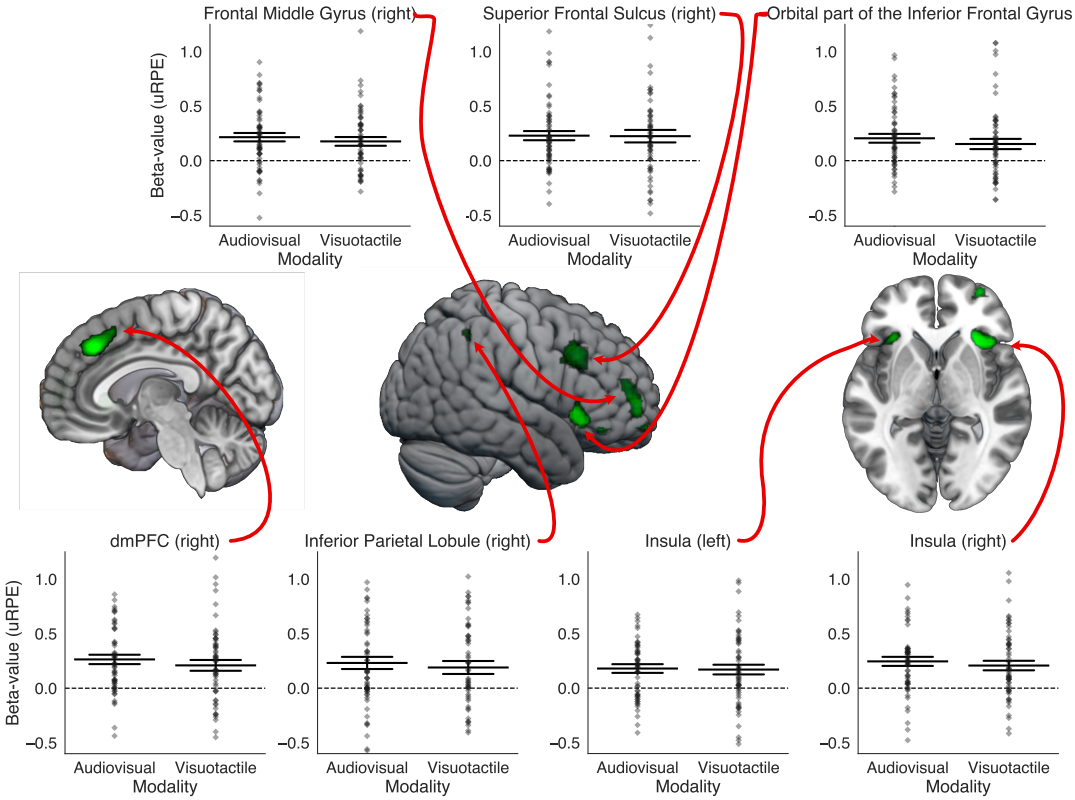

Figure 5: **Neural correlates of unsigned reward prediction errors (uRPEs)**. The brain images in the center display regions where unsigned RPEs, modeled as a parametric modulator at feedback onset, positively correlated with BOLD activation. uRPEs capture the magnitude of outcome surprise regardless of feedback valence. Statistical maps were thresholded at $t > 4.0$ and then $p < .05$, FWE-corrected using SnPM, and rendered on the MNI152 standard brain. Significant activations were observed in the right inferior, middle and superior frontal cortex, right dorso-medial prefrontal cortex (dmPFC), right inferior partial lobe, and bilateral insula. The accompanying strip plots display the mean beta estimates of the RPE contrast across all voxels within the thresholded cluster in MNI space, with each diamond marker representing data from an individual participant. Estimates are shown separately for audiovisual and visuotactile learning, revealing that RPE-related responses did not differ significantly between modalities. The widest horizontal line indicates the group mean, whereas the more narrow lines indicate a distance of one standard error of the mean (SEM).

To further test the dissociability of learning signals, we modeled uRPE, a valence-independent measure of outcome surprise computed at feedback. Conceptually, uRPE bridges RL and SL: like RPE, it depends on rewards, and like Shannon surprise, it reflects deviation from reward expectation. Given this conceptual overlap, we examined whether uRPE would recruit a distinct or shared network relative to RPE and surprise signals. In line with the former option, neural responses to uRPEs revealed a third, largely distinct network, with significant activation in the right dorsomedial PFC, bilateral insula, and a lateral frontoparietal circuit including inferior parietal lobule, middle frontal gyrus, orbital inferior frontal gyrus, and superior frontal sulcus (see Fig 5 and Table 3). This network aligns in part with prior reports of uRPE-related activation in the insula, dmPFC, and inferior parietal regions [37, 39], supporting their role in tracking reward salience and

behavioral relevance. However, compared to any previous findings, the frontal clusters identified in our study appear to be more lateral and ventral than any previously observed activations. This spatial shift may reflect the additional computational demands of our multisensory task, in which feedback integration and prediction updating occur across, rather than within, sensory modalities [6]. In such contexts, the brain must not only evaluate prediction errors but also determine which sensory stream contributed to the unexpected outcome and how these streams jointly inform future expectations. The recruitment of lateral and ventral frontal regions could therefore indicate enhanced engagement of cognitive control [56] and multisensory integration processes [57] needed to arbitrate between converging sensory inputs during outcome evaluation. Together, these results support the idea that uRPE engages a third, largely dissociable network, which is functionally distinct from both RPE- and surprise-related circuits, even when operating within the same multisensory learning task.

In a second set of whole-brain GLM analyses, we tested whether RL and SL are supported by modality-general or modality-specific mechanisms. We computed whole-brain contrasts for each learning signal (RPE, Shannon surprise, and uRPE), comparing visuotactile and audiovisual conditions (visuotactile > audiovisual and vice versa). No significant clusters emerged for any of the three learning signals. This suggests that, at the functional and anatomical resolution of our data, there is no evidence for modality-specific coding, neither in terms of increased or decreased activity in the identified regions, nor in the spatial distribution of the regions coding for these signals. Given the absence of significant modality-specific effects in the whole-brain analysis, we proceeded to examine whether the magnitude of learning signal coding in the already-identified domain-general learning ROIs differed across modalities. To this end, we extracted beta weights from the significant ROIs identified in the main contrasts (i.e., all voxels within clusters surviving the cluster-forming threshold). These beta weights showed comparable activation levels across both multisensory modalities (see swarm plots in Fig. 3, Fig. 4, and Fig. 5). In sum, these results suggest that the brain uses shared, domain-general systems to encode RPEs, surprise, and outcome uncertainty for multisensory learning, even when the precise sensory input channels differ, highlighting the generality and contextual flexibility of learning circuits in multisensory environments.

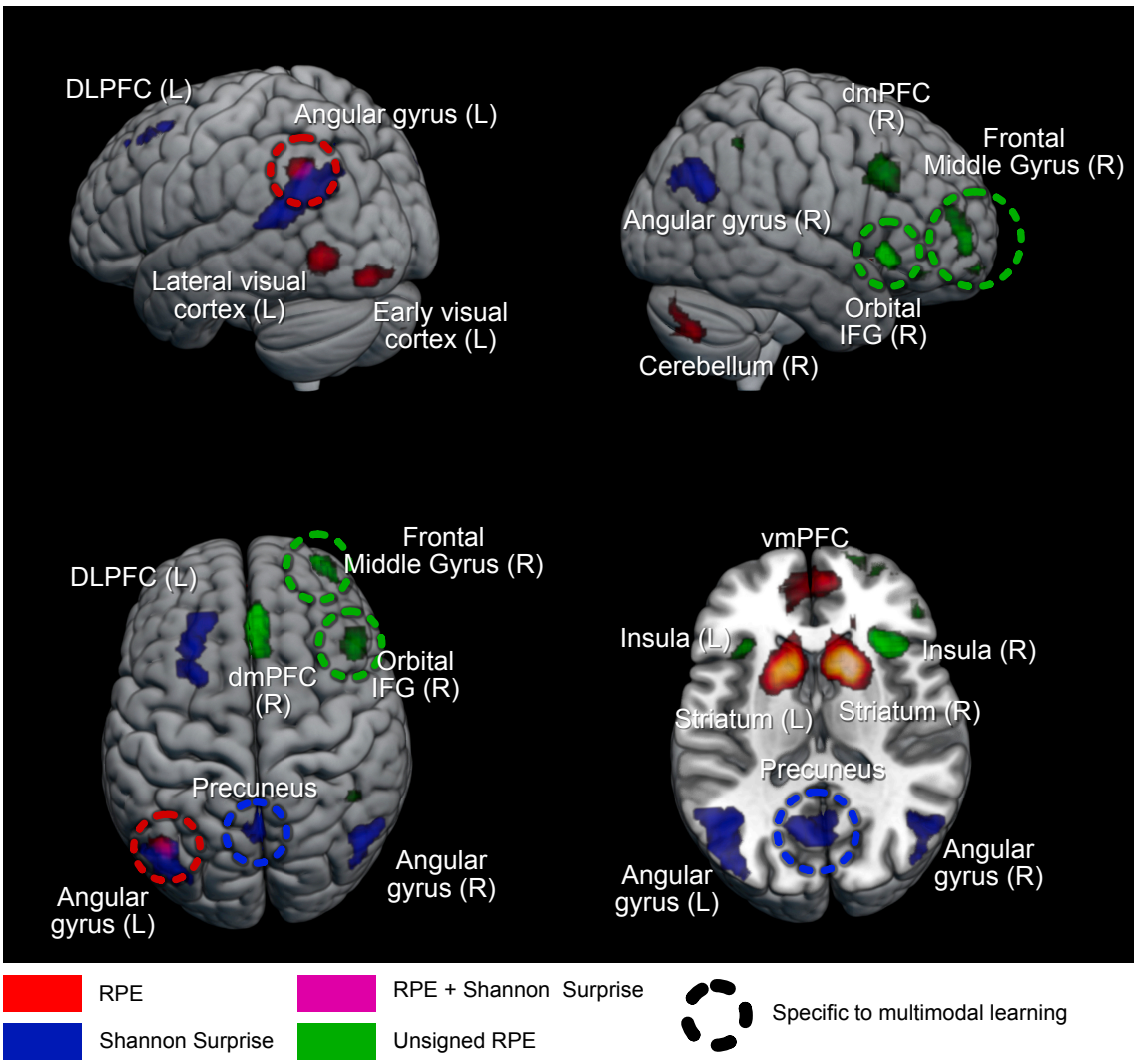

Figure 6: **Dissociable neural networks underlying multisensory learning.** Distinct brain regions encode reward prediction errors (RPE, red), Shannon surprise (blue), and unsigned reward prediction errors (uRPE, light green) during multisensory learning. The three networks show minimal overlap, highlighting specialized yet complementary systems for reinforcement learning, statistical learning, and outcome surprise. Regions outlined with a dotted line are not typically reported in unisensory learning tasks, suggesting unique contributions to multisensory integration. Brain maps are rendered on the MNI152 standard brain template.

Finally, we explored whether any regions may be specifically involved in multisensory learning, beyond their role in general RL and SL (Fig. 6). While most brain regions revealed in our GLM analyses overlap with known SL and RL networks [39, 53, 38, 37, 34, 35, 36, 37], several regions stood out (indicated by a dotted outlines in Fig. 6). Most notably, despite the largely distinct regions underlying RPE and Shannon surprise, the **left angular gyrus** consistently tracked both these signals. Prior studies have reported neighboring parietal regions like the inferior and superior parietal lobules for surprise [35, 34, 36], but have not implicated the angular gyrus as a locus for statistical learning [37, 39, 53, 38]. Similarly, while the inferior parietal lobule has been associated with RPEs in an extensive meta-analysis [37], the clusters from our findings and the meta-analysis do not overlap. In contrast, our findings show robust angular gyrus engagement in both statistical and reinforcement based multisensory learning, suggesting that this region (which sits at the intersection of the higher order visual, auditory, and tactile cortices) [58] may be a crucial hub for both types of learning about the associations between stimuli in these multiple senses.

We also observed surprise-related activity in the precuneus, a region widely recognized as a major cortical hub that integrates information from multiple sensory and cognitive networks [59, 60, 61]. This was despite the precuneus not being reported in prior SL studies [35, 34, 36, 37]. This suggests that the precuneus may also provide a modality-independent estimate of multisensory surprise. Regarding uRPE, several regions we observed (bilateral insula, inferior

parietal lobule, superior frontal sulcus, and dmPFC) have been linked to uRPEs before [37, 39]. However, the clusters in the frontal middle gyrus and orbital inferior frontal gyrus do not overlap with regions described in previous literature. Furthermore, regions typically associated with uRPEs in unisensory contexts - such as the striatum, cingulate cortex, and temporal regions - were not found to be reliably engaged in our study. Our results thus reveal distinct networks that underlie the three types of multisensory learning, which each comprise areas also implicated in unisensory learning but crucially also areas that appear uniquely recruited by learning about multisensory associations.

Discussion

Everyday learning often depends on non-redundant information distributed across multiple senses (e.g., the integration of taste and smell in perceiving flavor), yet most work has characterized learning with unisensory stimuli. Existing studies have modeled reinforcement learning with signed reward prediction errors [40, 62] that have been shown to drive learning behavior and neural activity in structures like ventral striatum/vmPFC [50, 47, 51, 52, 63]. In parallel, stimulus-structure-driven SL is often modeled as Shannon surprise and has been shown to drive both behavior and cortical responses in largely unisensory paradigms [64, 34, 35, 36]. To move beyond unisensory settings and identify learning about multi-sensory associations, as typical in everyday life, we designed a task in which informative structure existed only in multisensory stimulus combinations, not in the individual senses. Our novel task design allowed temporal and informational dissociation between SL - as indexed by stimulus-locked Shannon surprise - and RL indexed by feedback-locked RPE. Beyond the statistics and rewards introduced by our task design, participants' evolving expectations of reward outcomes motivated a third, feedback locked signal: uRPEs, which denotes outcome surprise and lies at the computational intersection of RL and SL. uRPEs share RL's dependence on expected rewards and SL's emphasis on deviation from these expectations, and are also known to modulate both learning behavior and neural responses in unisensory paradigms [65, 39, 37]. By isolating stimulus-driven structure, outcome-driven learning, and performance-contingent outcome surprise, our task and model-based fMRI results [46] thus provide a principled framework to study how the brain supports multiple forms of multisensory learning within a unified context.

Behaviorally, we observed the expected signatures of SL and RL. Response times increased with trial-wise Shannon surprise (hallmark of SL), indicating higher response times for rarer multisensory stimuli, and accuracy improved with reward feedback [34, 51]. Neurally, model-based fMRI revealed largely non-overlapping systems for each signal. RPEs recruited classical reward-related regions (ventral striatum, vmPFC) [47, 52, 63, 37, 38, 39], while Shannon surprise engaged a separate network in parietal and frontal cortices [36, 35]. uRPEs were encoded in a third, dissociable network involving the insula, dorsomedial prefrontal cortex, and frontoparietal regions [39, 37]. These brain regions overlap with both the salience network and the frontoparietal network, indicating that uRPEs may engage systems responsible for detecting behaviorally relevant events and implementing adaptive cognitive control [66, 67]. This suggests that the brain not only learns from environmental structure and feedback, but also monitors surprising feedback to potentially guide attentional shifts or cognitive control. Despite their distinct circuitry, all three signal (RPE, Shannon surprise, and uRPE) were encoded in modality-general networks, with no differences across audiovisual and visuotactile conditions. These findings suggest that the brain deploys specialized but flexible learning systems to support complex learning from multisensory input across various multi-sensory contexts. Nonetheless, modality-specific anatomical organization may exist at a finer scale than resolved by our standard-space fMRI analysis. Future work using high-resolution methods, such as electrophysiological recordings or participant-specific dense sampling [68, 69], could further elucidate this possibility.

Among the brain regions engaged by our task, the left angular gyrus stood out as a potential hub for multisensory learning. Despite the clear dissociation between SL and RL signals at the whole-brain level, this region consistently tracked both RPE and Shannon surprise, which are two computationally and temporally distinct signals, suggesting a role in integrating structure- and value-based learning signals across modalities. While neighboring parietal areas have previously been linked to either SL or RL in unisensory contexts [35, 36], the angular gyrus itself is not typically highlighted in meta-analyses of either computation. Its involvement here raises the possibility that it supports higher-order learning when both structure and feedback are distributed across multiple senses, consistent with proposals that angular gyrus functions as a multimodal integration hub [58]. This interpretation is supported by anatomical and neuropsychological evidence that the

angular gyrus is densely connected with higher-order visual, auditory, and somatosensory association cortices, making it well positioned to integrate crossmodal information [58, 70, 71]. Within predictive-coding frameworks, prediction errors are thought to be expressed in the cortical areas where the relevant information resides, enabling local updating of synaptic weights linking sensory inputs [72]. The emergence of prediction-error signals in the angular gyrus may therefore reflect the updating of multisensory associative links between modality-specific features. In large-scale network terms, the angular gyrus forms part of the salience and frontoparietal control networks, which are typically engaged in detecting and reorienting to behaviorally relevant sensory events [58]. These converging properties make it a plausible cortical locus for externally driven multisensory learning and integration. From a clinical perspective, the angular gyrus has been linked to the processing of word meaning [73, 74], and some studies have reported hypo-activation in this region [75, 76, 77, 78]. However, meta-analyses did not link angular gyrus activation specifically to dyslexia [79, 80]. While the angular gyrus regions we identified here may thus not be the main locus for the functions impaired in dyslexia, it contributes to integrating perceived information (written or spoken words) to their semantic meanings (abstract concepts), which is important for learning to read and write. In any case, our findings suggests that the angular gyrus supports higher-order multisensory learning and integration processes and may therefore be a crucial area for any cognitive function relying on this ability. This possibility may be tested by future studies using non-invasive brain stimulation techniques like TMS in conjunction with the paradigm and approach we introduced here.

We also observed surprise-related activity in the **precuneus**, a region that is densely interconnected with parietal, temporal, and prefrontal association cortices and is often proposed to be a cortical hub that integrates information from multiple sensory and cognitive networks [60, 61, 59]. Interestingly, this regions has not yet been linked to Shannon surprise [35, 34, 36, 37]. The presence of surprise-related signals in the precuneus may reflect the updating of internal, higher-order models that link multisensory inputs with self-referential or contextual expectations according to predictive coding frameworks [72]. In large-scale network terms, the precuneus is a core node of the default mode network, which supports internally oriented inference, mental model updating, and self-referential evaluation [59, 60, 81]. Its recruitment here thus complements the engagement of the angular gyrus, which is more closely associated with the salience and control networks and externally directed attention. Together, these patterns suggest that multisensory learning engages coordinated computations across networks supporting both externally triggered and internally oriented updating processes. In parallel, uRPE-related activation in frontal regions has not been consistently linked to feedback-driven surprise, unlike the lateral fronto-parietal cortex more commonly associated with the computation [39, 37]. Together, our findings suggest that multisensory learning may engage distinct or extended neural circuits beyond those observed in unisensory contexts. Notably, atypical activation patterns in the regions identified here have also been reported in developmental disorders affecting learning and cognitive control. In both Attention Deficit Hyperactivity Disorder (ADHD) [82, 83] and Autism Spectrum Disorder (ASD) [84], hypoactivation in frontal regions has been linked to deficits in executive functioning and attentional control. In dyslexia, frontal activation patterns are more heterogeneous, with both hypo- and hyperactivation reported across studies [79, 85]. The observed hypoactivation has been interpreted as reflecting executive dysfunction [86] and difficulties in accessing lexical and sublexical phonological representations [79], and it has also been associated with the high comorbidity between dyslexia and ADHD [86]. Conversely, hyperactivation has been suggested to reflect articulatory compensatory mechanisms [79] or a broader spatial distribution of the reading network [87]. Of particular relevance, the inferior frontal gyrus, together with the posterior superior temporal sulcus, occipito-temporal cortex, and inferior parietal lobe, forms the core network underlying audiovisual letter–speech sound learning and integration [88]. Both children and adults with dyslexia exhibit reduced activation within this network, as well as impaired functional and structural connectivity among its constituent regions [88]. Atypical hyperactivation of the precuneus has likewise been observed in ASD [84, 89] and ADHD [89], likely reflecting insufficient deactivation of the default mode network. In dyslexia, by contrast, reduced precuneus activation relative to typical readers has been reported [85, 90], suggesting increased cognitive effort and compensatory recruitment during reading-related tasks [90, 80]. Taken together, these clinical findings highlight that the frontal and parietal regions engaged in uRPE processing during multisensory association learning are central to learning and control processes disrupted in developmental disorders.

As the first study to simultaneously dissociate and examine SL, RL, and their intersection in a multisensory context, our results provide novel insights into how the brain integrates structure and feedback across sensory modalities, and offer a strong rationale for future causal and developmental studies to elucidate the specific contributions of these regions to the emergence of learning disorders.

Several caveats warrant mention. First, as with all fMRI studies, our results are only correlational, so that causal approaches such as TMS [91, 92, 93] or lesion studies [94] are needed to test whether regions that covary in their activity with the learning signals are indeed necessary for multisensory learning. While superficial regions like the angular gyrus can be reached with TMS, deep or midline structures may be more effectively targeted using emerging methods such as focused ultrasound [95, 96] or deep-TMS [97]. Studying developmental and aging populations could also reveal how multisensory learning and the integration of value and structure evolve across the lifespan [98]. Finally, extending this framework to ecologically valid domains such as speech processing, reading, or sensorimotor coordination could reveal how the multisensory learning mechanisms identified here contribute to well-established cognitive systems. This appears promising since the precuneus and frontal regions identified here partially overlap with networks involved in letter-speech sound learning [88] and with regions implicated in developmental disorders such as dyslexia, ASD, and ADHD [89, 90, 82, 84]. Applying our modeling approach to these contexts would therefore help determine whether the same neural computations that integrate structure and feedback across modalities also support language and reading acquisition, and are impaired in people with atypical development. Together, these avenues offer promising paths for a better understanding of the computational and neural processes underlying multisensory learning and their possible contributions to a wide range of cognitive and neurodevelopmental disorders.

Methods

Participants

Sixty-four right-handed participants (21 women; age range 18–30 years, mean age = 23.3) took part in the fMRI study. All participants were screened for MRI compatibility prior to inclusion and reported no psychiatric or neurological disorders or need for visual correction. Participants were students at the University of Zurich or ETH Zurich, with those enrolled in Economics, Psychology, or Computer Science excluded from participation. Six participants were removed from the analyses due to scanner-related problems or excessive head motion, resulting in a final sample of 58. All procedures conformed to the Declaration of Helsinki and were approved by the Ethics Committee of the Canton of Zurich.

Procedure

The fMRI study was conducted at the Laboratory for Social and Neural Systems Research (SNS Lab) of the University Hospital Zurich. Each session lasted no more than 2 hours and 15 minutes. Upon arrival, participants completed MRI safety screening and provided written informed consent. They then proceeded to a behavioral testing room, where they read instructions for the multisensory learning task, answered comprehension questions, and received additional information on MRI safety. Participants subsequently performed the multisensory learning task inside the MRI scanner (Philips ACHIEVA 3T), during which both behavioral and neural measures were collected. Visual stimuli were projected onto a screen and viewed via a mirror mounted on the head coil. Auditory stimuli were delivered using MRI-compatible headphones (MR Confon GmbH, www.mr-confon.de), and tactile stimuli were applied to the left index finger using a piezoelectric stimulator (mini-PTS, Dancer Design, http://www.dancerdesign.co.uk/). Eye movements were recorded with an MR-compatible infrared Eyelink II CL v.4.51 eye-tracker (SR Research Ltd), and physiological data were obtained using a breathing belt and an MR-compatible ECG device. In addition to functional imaging, we acquired a high-resolution anatomical scan and diffusion tensor imaging (DTI) data. Participants received a base payment of 50 CHF, plus a performance-dependent bonus of up to 48 CHF.

Multisensory learning task

Participants were instructed to play a game in which they took the role of a scientist studying newly discovered insects. On each trial, a visual image of an insect was paired with either an auditory "call" (a sequence of beeps presented over headphones) or a tactile "dance" (a sequence of vibrations delivered to the left index finger). Participants judged whether the multisensory pairing would attract a mate ("attract" vs. "no attract") with a button press.

Visual stimuli. Each block featured three insect "species," distinguished solely by the orientation of a textured noise pattern superimposed on a butterfly-shaped silhouette (Fig. 1a). The

textures were generated by filtering pink noise with von Mises orientation filters, yielding three distinct orientation-defined patterns at approximately 22.5°, 67.5°, and 112.5°. These angles were chosen to be equally spaced away from the cardinal axes. Images were exported at high resolution (1200 dpi), presented centrally on a gray background. Importantly, species identity could be determined from this single visual feature (orientation).

Auditory and tactile stimuli. Auditory "calls" and tactile "dances" were constructed as sequences of discrete bursts presented within a fixed 1.5 s stimulus window. Each pattern contained 9 bursts: 8 tone (or vibration) bursts separated by variable inter-burst intervals, plus a final burst, such that the entire sequence always filled 1.5 s. Each burst lasted 15 ms. For auditory stimuli, bursts were sine tones (delivered via MRI-compatible headphones), while for tactile stimuli, bursts were square pulses delivered to the left index finger via a piezoelectric stimulator (mini-PTS, Dancer Design; 48 kHz sampling rate). For both modalities, three distinct temporal patterns were created per set by varying the distribution of inter-burst intervals while keeping total duration fixed. This ensured that tactile and auditory stimuli were structurally analogous and directly comparable across modalities.

Task structure and timing. The experiment consisted of six blocks (runs), alternating between audiovisual (calls) and visuotactile (dances) modalities (3 blocks each). The starting modality was randomized across participants. Each block contained 60 trials. Trial timing was as follows: the multisensory stimulus was presented for 1.5 s (during which responses were recorded), followed by a jittered inter-stimulus interval (0.5–1.5 s), then feedback displayed for 1 s. Inter-trial intervals were jittered between 3–5 s to optimize design efficiency and minimize correlations between parametric regressors. A response-time window of up to 3.0 s was enforced (responses after this window were treated as missing). The total task duration was approximately 54 minutes.

Reinforcement-learning (feedback) contingencies and calibration. Each insect species had exactly one "high-success" pairing (i.e., one specific call or dance most likely to attract a mate), whereas the other two pairings for that species had much lower success rates. Feedback was probabilistic: correct pairings led to "+1" (green tick) most of the time, while incorrect pairings yielded "0" (red cross), with occasional probabilistic flips to maintain uncertainty (participants were informed that feedback was probabilistic but were not given exact probabilities). The first two blocks served as calibration (feedback probability fixed at 0.8 for correct responses). Based on each participant's calibration accuracy, subsequent blocks used individually adjusted correct-feedback probabilities of 0.7, 0.8, or 0.9 (thresholds: $< 0.56 \Rightarrow 0.9$; 0.56–$< 0.70 \Rightarrow 0.8$; $\geq 0.70 \Rightarrow 0.7$), ensuring sustained learning pressure across the session. Reward magnitude was fixed at 1.0 "point" for each correct outcome. These points were then converted into monetary payment according to performance-dependent scaling factors. Specifically, participants started with a base rate of 50 CHF, and their accumulated points were translated into an additional bonus ranging from 0 to 48 CHF. During the two initial calibration blocks, reward probabilities were fixed (0.8 for correct feedback) and a default conversion factor was used. Based on accuracy in these blocks, participants were assigned to one of three difficulty levels (easy, medium, hard) for the subsequent blocks, which determined both the probabilistic structure of feedback (0.9, 0.8, or 0.7 probability of positive feedback) and the money conversion factor (with harder conditions yielding higher monetary weightings). This ensured that overall payment opportunities were balanced across individuals while keeping the task suitably challenging.

Statistical-learning (stimulus frequency) structure. In addition to the reinforcement structure, we embedded a statistical structure in the frequencies of the nine possible multisensory combinations (3 visual species $\times$ 3 calls/dances). Frequencies followed a 3×3 "joint probability" matrix with three levels: **common** ($p = 0.50$ of the trials for that visual species), **occasional** ($p = 0.35$), and **rare** ($p = 0.15$). Specifically, the three common pairs were $(0, A)$, $(1, B)$, and $(2, C)$; the three occasional pairs were $(0, B)$, $(1, C)$, and $(2, A)$; and the three rare pairs were $(0, C)$, $(1, A)$, and $(2, B)$.

Trial sequences were generated algorithmically to meet three constraints: (1) within every 20-trial segment of a block, pair counts matched the expected ratios exactly (10 common, 7 occasional, 3 rare events); (2) across the full 60 trials of a block, frequencies matched the joint-probability matrix precisely; and (3) immediate repetitions in either unisensory modality were avoided (e.g., if the current trial was visual "1" with call "A," the following trial could not share the same visual or the same auditory/tactile element).

These constraints ensured that the statistical structure was both robust and evenly distributed, while preventing trivial strategies such as detecting simple repetitions. Importantly, participants were never informed about the frequency manipulation, and the statistical structure was independent of reward contingencies, allowing statistical learning to be probed implicitly.

Computational models

Reinforcement Learning

To model reinforcement learning computations that are explicitly required for task performance, we employed Q-learning with two 3×3 Q-value tables. Each table represented the Q-values for the two possible actions (attract or not attract) across the nine multisensory stimuli. This setup allowed us to estimate the value of each action for each stimulus pairing, as shown in Figure 1 from the main text. The green cells in the figure represent the pairings most likely to attract a partner (and therefore requiring a "yes" response; please note that rewards were (or were not) given for both correct (or incorrect) " yes" and "no" responses, so that rewards were independent of statistical structure and response type). Q-learning assumes that participants update their Q-values based on the RPE which is the difference between the expected value and the observed reward. Unless otherwise specified, unchosen values remained unchanged. In addition to signed RPEs, we also derived uRPEs, defined as the absolute magnitude of the RPE, indexing the degree of outcome surprise irrespective of valence [39, 37].

We tested different versions of Q-learning to capture the diverse strategies and biases that participants might have. These models included:

1. **Basic Q-learning** — The simplest model was the basic Q-learning algorithm where participants only updated the value for the chosen response, with one learning rate α for all possibilities [40, 20].
$$Q_c(t) = Q_c(t-1) + \alpha\big(R - Q_c(t-1)\big)$$
where Q_c denotes the value of the chosen option for the presented multisensory state in that trial. The Q-values were initialized to 0.5 for both actions in this and all other models, except for versions of the Vinit model where initial values were treated as free parameters.

2. **Asym** — Studies have suggested that people learn differently from positive and negative RPEs [41, 42, 43, 44]. To incorporate this, we designed a model with separate learning rates for positive and negative feedback:
$$Q_c(t) = Q_c(t-1) + \alpha_{pos}\big(R - Q_c(t-1)\big), \quad \text{if } (R - Q_c(t-1)) \geq 0,$$
$$Q_c(t) = Q_c(t-1) + \alpha_{neg}\big(R - Q_c(t-1)\big), \quad \text{if } (R - Q_c(t-1)) < 0.$$

3. **Transfer** — Given the task structure, each trial conveyed information not only about the presented multisensory stimulus but also about related stimuli that shared one unisensory feature. For example, if A1 attracted a partner, then A2, A3, B1, and C1 could be inferred to be less likely. To model this inference, we allowed related, unchosen pairings to be updated with a scaled parameter:
$$Q_c^o(t) = Q_c^o(t-1) - \kappa\alpha\big(R - Q_c(t)\big),$$
where Q_c^o denotes the value of a related option, and $\kappa \leq 1$ scales the transfer update so that related stimuli are never updated more strongly than the chosen stimulus.

4. **Vinit** — This model captured initial biases in action preferences. Instead of fixing Q-values at 0.5, the initial Q-values for both action tables were treated as free parameters, allowing us to model a priori action propensities. The free parameters were Q_0, Q_1, α. Updating followed the same rule as in basic Q-learning.

5. **TransferVinit** — This model combined the Transfer and Vinit features, i.e., transfer updates were applied and initial Q-values were free parameters.

All of the above models had parameter recovery correlations above 0.6 between simulated and recovered parameters. Therefore, their parameters were considered recoverable. We also explored additional variants, such as models with asymmetric learning rates for actions, Pearce–Hall–style adaptive learning, and dynamic learning rates, but these failed to reach the recovery threshold ($r < 0.6$) and were excluded from further fitting.

The unrecoverable models included:

- **Extra** — Asymmetric learning rates for the two different actions:

$$Q_{attract}(t+1) = Q_{attract}(t) + \alpha_{attract}\big(R - Q_{attract}(t)\big),$$

$$Q_{\neg attract}(t+1) = Q_{\neg attract}(t) + \alpha_{\neg attract}\big(R - Q_{\neg attract}(t)\big).$$

- **Pearce** — A Pearce–Hall style model [65] with trial-wise adaptive learning rates:

$$\alpha_t = \eta + \omega\,|R - Q_{chosen}(t)|,$$

  where η is the base learning rate and ω modulates sensitivity to prediction errors.

- **Dyna** — A dynamic learning-rate model:

$$\alpha_t = \alpha_{t-1} + \omega\big(|R - Q_{chosen}(t)| - \alpha_{t-1}\big).$$

In the end, only the five models with recoverable parameters (Basic, Asym, Transfer, Vinit, TransferVinit) were fitted. All models were fit to each participant's trial-by-trial choices via maximum likelihood (SciPy v1.10; L-BFGS-B; multiple random initializations). Model evidence was approximated with the Bayesian Information Criterion (BIC), computed from the maximum log-likelihood and number of free parameters. For model selection, we summed BIC across the six blocks per participant and chose the lowest-BIC model as that participant's winner. Of the five recoverable models, three accounted for all participants' data: **Basic Q-learning (n = 31), Transfer (n = 16), and Asymmetric learning rates (n = 11)**; **no participants** were best fit by **Vinit** or **Transfer+Vinit**. The per-participant winning model was then used to generate trial-wise value estimates, signed RPEs, and uRPEs for all behavioral and fMRI analyses.

Statistical Learning

To model statistical learning, we leveraged the probabilistic structure of the task. Each visual species could be paired with one of three calls/dances, resulting in nine possible multisensory combinations. Frequencies followed a 3×3 joint-probability matrix with three levels: common ($p = 0.50$), occasional ($p = 0.35$), and rare ($p = 0.15$). This structure yielded a 50/50 split between pairings that were most likely versus least likely to attract a partner, preventing participants from adopting a trivial strategy of selecting the same action throughout the experiment.

We modeled statistical learning using a Bayesian updating framework without any free parameters, following [33]. The model assumes a uniform prior over the nine possible stimulus combinations at the start of each block. On every trial, beliefs about the probability of encountering each stimulus i are updated, resulting in a Dirichlet posterior that gradually converges toward the true empirical frequencies.

On each trial t, the belief for stimulus i is given by:

$$p_i(t) = \frac{1 + \sum_{t'=1}^{t} I_{[o(t')=i]}}{9 + t},$$

where $I_{[o(t')=i]}$ is an indicator function equal to 1 if stimulus i was observed on trial t', and 0 otherwise. This formulation reflects Bayesian updating with a uniform prior and counts-based likelihood.

To quantify statistical learning, we computed the Shannon surprise of the observed stimulus o on trial t:

$$\text{Surprise}_o(t) = -\log\big(p_o(t)\big).$$

Lower prior belief $p_o(t)$ results in higher surprise. Trial-wise surprise values were then used as a proxy for statistical learning, with the hypothesis that more surprising stimuli would elicit longer response times.

Behavioral analysis

To assess how reinforcement and statistical learning signals influenced behavior, we fit a linear mixed-effects regression model predicting trial-wise response times:

$$\text{responseTime} \sim 1 + \text{surprise}_o + Q_c + Q_u + (1\,|\,\text{subject}),$$

where surprise$_o$ is the Shannon surprise of the presented stimulus on trial t, Q_c and Q_u are the chosen and unchosen action values from the fitted RL model, and $(1 \mid \text{subject})$ specifies subject-level random intercepts.

All predictors were z-scored prior to analysis to enable direct comparison of regression coefficients. Mixed-effects models were estimated in `R` using the `lmerTest` package, called from Python via `rpy2`. Statistical significance was evaluated using Satterthwaite's approximation for degrees of freedom as implemented in `lmerTest`.

MRI acquisition and preprocessing

Functional and anatomical MRI data were acquired at the Laboratory for Social and Neural Systems Research, University Hospital Zurich, using a Philips Achieva 3T whole-body MR scanner equipped with a 32-channel head coil. Functional images were collected across six runs using a T2*-weighted gradient-recalled echo-planar imaging (EPI) sequence (222 volumes + 5 dummy scans; flip angle = 90°; repetition time, TR = 2334 ms; echo time, TE = 30 ms; matrix size = 80 × 78; field of view = 240 × 240 mm; in-plane resolution = 3 mm; 40 axial slices, slice thickness = 3 mm, slice gap = 0.5 mm; SENSE acceleration factor = 1.5 in the phase-encoding direction; acquisition time = 8:54 min).

A high-resolution T1-weighted structural image was acquired using an MPRAGE sequence (field of view = 256 × 256 × 170 mm; 1 mm isotropic resolution; inversion time, TI = 2800 ms; 170 sagittal slices; flip angle = 8°; TR = 8.3 ms; TE = 3.9 ms; SENSE acceleration factor = 2 in the left–right direction; acquisition time = 5:35 min). In addition, diffusion tensor imaging (DTI) data were acquired (field of view = 224 × 224 × 118 mm; 2 mm isotropic resolution; 59 axial slices; TR = 9000 ms; 64 non-collinear diffusion-weighted directions with b = 1000 s/mm^2, plus one non-diffusion-weighted image with b = 0 s/mm^2; acquisition time = 10:03 min).

Preprocessing of the fMRI data was performed using *fMRIPrep* 20.2.3 [99] (RRID:SCR_016216), which is based on Nipype 1.6.1 [100] (RRID:SCR_002502) and incorporates routines from SPM. The pipeline included realignment and unwarping, slice-timing correction, coregistration, segmentation, and normalization. A full description of preprocessing steps is provided in the Appendix. Finally, spatial smoothing was applied using SPM12, with a 6 mm full-width-at-half-maximum (FWHM) Gaussian kernel to reduce noise and account for residual inter-individual differences in anatomy during group-level analyses.

First-Level Analysis

To model task-related neural activity, we specified a general linear model (GLM) in SPM12 (build 7771) implemented in MATLAB (R2023b, version 23.2.0.2668659). The GLM was designed to capture blood-oxygen-level-dependent (BOLD) responses associated with audiovisual and visuotactile statistical and reinforcement learning. The design matrix included six motion parameters (three translation and three rotation estimates) to account for head motion artifacts, the first five aCompCor components derived from fMRIPrep to model noise related to physiological and non-neural signals, and physiological noise regressors (cardiac and respiratory signals) estimated using the PhysIO Toolbox for MATLAB.

The task-related regressors included the choice period, modeled as an epoch with a duration equal to the participant's response time and parametrically modulated by Shannon surprise and the value of the chosen option. The feedback period was modeled as a fixed-duration epoch (1 second) parametrically modulated by RPE and uRPE. All parametric modulators were z-standardized (mean-centered and scaled to unit variance) to ensure comparability across participants and conditions. No orthogonalization was applied to the regressors of interest to preserve shared variance between predictors. To verify that these regressors captured distinct computational signals, we quantified their trial-wise correlations across participants. Mean Pearson correlations were very small, confirming minimal collinearity. Mean Pearson correlations were small, confirming minimal collinearity between the learning signals. Across subjects, mean coefficients of determination were RPE–SPE, $r^2 = 0.007 \pm 0.007$; uRPE–SPE, $r^2 = 0.045 \pm 0.027$; and RPE–uRPE, $r^2 = 0.110 \pm 0.118$. These low shared variances indicate that the three learning signals were largely independent and suitable for simultaneous inclusion as separate regressors in the fMRI general linear model (GLM). The GLM design further accounted for any residual shared variance among regressors.

For each parametric modulator, we defined first-level contrasts for all runs vs. implicit baseline, audio runs vs. baseline, tactile runs vs. baseline, and audio runs vs. tactile runs.

Group Analysis

For second-level inference, we employed non-parametric permutation testing using the SnPM toolbox (v13.01). This approach involved 5,000 permutations to generate a null distribution for statistical testing, cluster-wise inference with family-wise error (FWE) correction at $p < 0.05$ to control for multiple comparisons. We tested both positive and negative contrasts for RPE, Shannon surprise, and uRPE. To assess modality-specific effects, we also computed direct contrasts between audiovisual and visuotactile conditions for each variable of interest.

Cluster-forming thresholds were optimized for each analysis to balance cluster overlap and statistical sensitivity, with cluster-forming thresholds set at $t > 8.0$ for RPE, $t > 3.1$ for Shannon surprise, and $t > 4.0$ for uRPE.

ROI Analysis

For the region-of-interest (ROI) analysis, we defined each ROI as the set of all voxels in MNI space that survived cluster-wise family-wise error (FWE) correction (at p ¡ 0.05) and exceeded the variable-specific cluster-forming threshold (see previous section). These voxels, identified in the second-level group analysis, formed the functional masks for each ROI. From each mask, we extracted the mean beta estimate per participant for the audiovisual and visuotactile regressors, focusing on the contrasts of interest: RPE, Shannon surprise, and uRPE. By comparing these mean modality-specific beta estimates across conditions, we directly tested for differences in neural activation patterns related to RPE, surprise, and uRPE between audiovisual and visuotactile learning, while accounting for multiple comparisons through cluster-based thresholding.

A fMRIPrep

The preprocessing steps for the anatomical image were as follows:

A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset.The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al. 2010), distributed with ANTs 2.3.3 (Avants et al. 2008, RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823, Zhang, Brady, and Smith 2001). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847, Dale, Fischl, and Sereno 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438, Klein et al. 2017). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c [Fonov et al. (2009), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],

For each of the 6 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. A B0-nonuniformity map (or fieldmap) was estimated based on a phase-difference map calculated with a dual-echo GRE (gradient-recall echo) sequence, processed with a custom workflow of SDCFlows inspired by the epidewarp.fsl script and further improvements in HCP Pipelines (Glasser et al. 2013). The fieldmap was then co-registered to the target EPI (echo-planar imaging) reference run and converted to a displacements field map (amenable to registration tools such as ANTs) with FSL's fugue and other SDCflows tools. Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve and Fischl 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, Jenkinson et al. 2002). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 (Cox and Hyde 1997, RRID:SCR_005927). The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): fsaverage, fsnative. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by Power et al. 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor, Behzadi et al. 2007). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's asegsegmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original

implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al. 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

A.1 Contrast 1: RPE

Table 1: Clusters that correlate with RPE, cluster-forming threshold T = 8.

Cluster Label	**Voxels**	**MAX T**	**MAX X (mm)**	**MAX Y (mm)**	**MAX Z (mm)**	**P FWE-CORR.**
Caudate, Putamen R	391	19.29	18	14	-12	0.0002
Caudate, Putamen L	273	16.22	-16	14	-12	0.0002
Angular Gyrus L	66	11.78	-46	-66	44	0.0002
Visual cortex (Mid, Lat Occipital) L	139	11.78	-46	-82	2	0.0002
vmPFC	213	11.05	8	54	-8	0.0002
Visual Cortex (Lingual, Fusiform) L	61	10.80	-22	-88	-12	0.0002
Cerebellum R	111	10.67	42	-70	-40	0.0002

A.2 Contrast 5: Surprise

Table 2: Clusters that correlate with surprise, cluster-forming threshold T = 3.1.

Cluster Label	**Voxels**	**MAX T**	**MAX X (mm)**	**MAX Y (mm)**	**MAX Z (mm)**	**P FWE-CORR.**
Precuneus L	201	6.01	-4	-54	40	0.0094
DLPFC L	103	4.98	-28	24	48	0.0284
Angular Gyrus R	90	4.83	48	-72	37	0.0354
Angular gyrus L	111	4.68	-36	-82	40	0.0082

A.3 RL negative T-value neural signatures

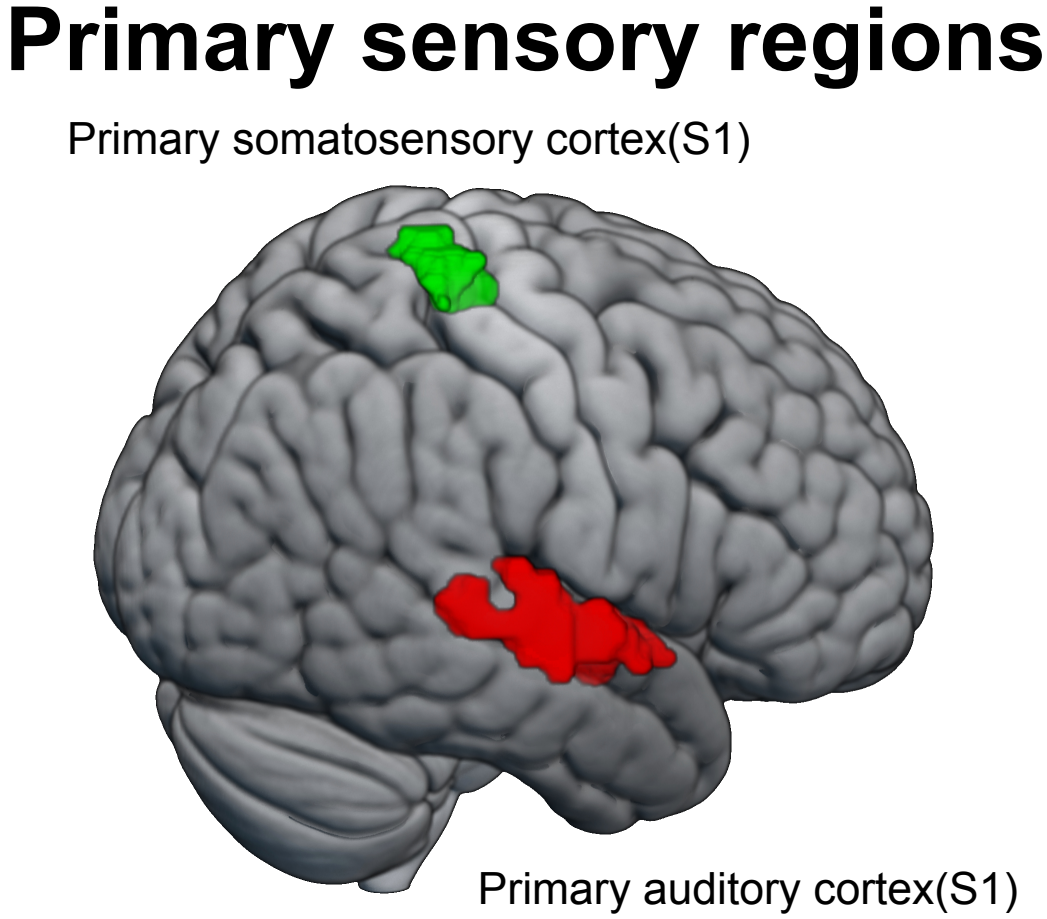

Figure S1: Primary sensory regions that are found with the contrast $\text{choice}_{\text{audiovisual}} > \text{choice}_{\text{visuotactile}}$, thresholded at t=3.1

A.4 Unsigned RPE

Table 3: Clusters that correlate with unsigned RPE (URPE), cluster-forming threshold T = 4.0

Cluster Label	**Voxels**	**MAX T**	**MAX X (mm)**	**MAX Y (mm)**	**MAX Z (mm)**	**P FWE-CORR.**
dmPFC R	142	8.02	2	32	37	0.0010
Insula R	116	7.25	36	24	-5	0.0012
Insula L	31	5.47	-34	26	-2	0.0088
Frontal Middle Gyrus R	123	4.68	32	60	2	0.010
Orbital part of the Inferior Frontal Gyrus R	13	5.32	50	36	-12	0.032
Superior Frontal Sulcus R	126	5.05	48	24	37	0.0010
Inferior Parietal Lobule R	24	xxx	44	-46	48	0.0110

References

[1] Gemma Calvert, Charles Spence, and Barry E Stein. *The handbook of multisensory processes*. MIT press, 2004.

[2] Shinsuke Shimojo and Ladan Shams. "Sensory modalities are not separate modalities: plasticity and interactions". In: *Current Opinion in Neurobiology* 11.4 (2001), pp. 505–509. ISSN: 0959-4388. DOI: https://doi.org/10.1016/S0959-4388(00)00241-5. URL: https://www.sciencedirect.com/science/article/pii/S0959438800002415.

[3] Asif A Ghazanfar and Charles E Schroeder. "Is neocortex essentially multisensory?" In: *Trends in cognitive sciences* 10.6 (2006), pp. 278–285.

[4] E Nava, M Giraud, and N Bolognini. *The emergence of the multisensory brain: From the womb to the first steps. iScience, 27 (1), 108758*. 2023.

[5] Qihao Zheng and Yong Gu. "From Multisensory Integration to Multisensory Decision-Making". In: *Advances of Multisensory Integration in the Brain* (2024), pp. 23–35.

[6] Jon Driver and Toemme Noesselt. "Multisensory interplay reveals crossmodal influences on 'sensory-specific'brain regions, neural responses, and judgments". In: *Neuron* 57.1 (2008), pp. 11–23.

[7] J Vroomen et al. "The Handbook of Multisensory Processes". In: (2004).

[8] Barry E Stein and Terrence R Stanford. "Multisensory integration: current issues from the perspective of the single neuron". In: *Nature reviews neuroscience* 9.4 (2008), pp. 255–266.

[9] Micah M Murray and Mark T Wallace. *The neural bases of multisensory processes*. CRC press, 2011.

[10] Solange Denervaud et al. "Multisensory Gains in Simple Detection Predict Global Cognition in Schoolchildren". In: *Scientific Reports* 10.1 (Feb. 4, 2020), p. 1394. ISSN: 2045-2322. DOI: 10.1038/s41598-020-58329-4. URL: https://www.nature.com/articles/s41598-020-58329-4 (visited on 07/30/2024).

[11] Emmanuelle Dionne-Dostie et al. "Multisensory Integration and Child Neurodevelopment". In: *Brain Sciences* 5.1 (Feb. 11, 2015), pp. 32–57. ISSN: 2076-3425. DOI: 10.3390/brainsci5010032. URL: https://www.mdpi.com/2076-3425/5/1/32 (visited on 07/03/2024).

[12] Gary K Beauchamp and Julie A Mennella. "Early flavor learning and its impact on later feeding behavior". In: *Journal of pediatric gastroenterology and nutrition* 48 (2009), S25–S30.

[13] Malika Auvray and Charles Spence. "The multisensory perception of flavor". In: *Consciousness and cognition* 17.3 (2008), pp. 1016–1031.

[14] Charles Spence. "Multisensory flavor perception". In: *Cell* 161.1 (2015), pp. 24–35.

[15] Nada Frei et al. "Toward a Mechanistic Understanding of Reading Difficulties: Deviant Audiovisual Learning Dynamics and Network Connectivity in Children with Poor Reading Skills". In: *Journal of Neuroscience* 45.17 (2025).

[16] Noemi Hahn, John J Foxe, and Sophie Molholm. "Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia". In: *Neuroscience & Biobehavioral Reviews* 47 (2014), pp. 384–392.

[17] Sally E Shaywitz et al. "Prevalence of reading disability in boys and girls: Results of the Connecticut Longitudinal Study". In: *Jama* 264.8 (1990), pp. 998–1002.

[18] Robert M Yerkes and Sergius Morgulis. "The method of Pawlow in animal psychology." In: *Psychological Bulletin* 6.8 (1909), p. 257.

[19] B Frederic Skinner. "Two types of conditioned reflex and a pseudo type". In: *The Journal of General Psychology* 12.1 (1935), pp. 66–77.

[20] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[21] Horace B Barlow et al. "Possible principles underlying the transformation of sensory messages". In: *Sensory communication* 1.01 (1961), pp. 217–233.

[22] Eero P Simoncelli and Bruno A Olshausen. "Natural image statistics and neural representation". In: *Annual review of neuroscience* 24.1 (2001), pp. 1193–1216.

[23] Wilson S Geisler. "Visual perception and the statistical properties of natural scenes". In: *Annu. Rev. Psychol.* 59.1 (2008), pp. 167–192.

[24] Edward C Tolman. "Cognitive maps in rats and men." In: *Psychological review* 55.4 (1948), p. 189.

[25] Matthew Botvinick et al. "Reinforcement learning, efficient coding, and the statistics of natural tasks". In: *Current opinion in behavioral sciences* 5 (2015), pp. 71–77.

[26] Alec Solway et al. "Optimal behavioral hierarchy". In: *PLoS computational biology* 10.8 (2014), e1003779.

[27] Lucy Lai, Ann Zixiang Huang, and Samuel J Gershman. "Action chunking as policy compression". In: (2022).

[28] Peter Dayan. "Improving generalization for temporal difference learning: The successor representation". In: *Neural computation* 5.4 (1993), pp. 613–624.

[29] Samuel J Gershman. "The successor representation: its computational logic and neural substrates". In: *Journal of Neuroscience* 38.33 (2018), pp. 7193–7200.

[30] C. E. Shannon. "A mathematical theory of communication". In: *The Bell System Technical Journal* 27.3 (1948), pp. 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[31] Aaron M Bornstein and Nathaniel D Daw. "Dissociating hippocampal and striatal contributions to sequential prediction learning". In: *European Journal of Neuroscience* 35.7 (2012), pp. 1011–1023.

[32] A Ross Otto et al. "The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive". In: *Psychological science* 24.5 (2013), pp. 751–761.

[33] Arkady Konovalov and Ian Krajbich. "Gaze data reveal distinct choice processes underlying model-based and model-free reinforcement learning". In: *Nature communications* 7.1 (2016), p. 12438.

[34] Arkady Konovalov and Ian Krajbich. "Neurocomputational dynamics of sequence learning". In: *Neuron* 98.6 (2018), pp. 1282–1293.

[35] Antonino Visalli et al. "Bayesian modeling of temporal expectations in the human brain". In: *NeuroImage* 202 (2019), p. 116097.

[36] Jill X O'Reilly et al. "Dissociable effects of surprise and model update in parietal and anterior cingulate cortex". In: *Proceedings of the National Academy of Sciences* 110.38 (2013), E3660–E3669.

[37] Philip R Corlett, Jessica A Mollick, and Hedy Kober. "Meta-analysis of human prediction error for incentives, perception, cognition, and action". In: *Neuropsychopharmacology* 47.7 (2022), pp. 1339–1349.

[38] Jane Garrison, Burak Erdeniz, and John Done. "Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies". In: *Neuroscience & Biobehavioral Reviews* 37.7 (2013), pp. 1297–1310.

[39] Elsa Fouragnan, Chris Retzler, and Marios G Philiastides. "Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis". In: *Human brain mapping* 39.7 (2018), pp. 2887–2906.

[40] Robert A Rescorla. "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement". In: *Classical conditioning, Current research and theory* 2 (1972), pp. 64–69.

[41] Nathaniel D Daw, Sham Kakade, and Peter Dayan. "Opponent interactions between serotonin and dopamine". In: *Neural networks* 15.4-6 (2002), pp. 603–616.

[42] Michael J Frank et al. "Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning". In: *Proceedings of the National Academy of Sciences* 104.41 (2007), pp. 16311–16316.

[43] Germain Lefebvre et al. "Behavioural and neural characterization of optimistic reinforcement learning". In: *Nature Human Behaviour* 1.4 (2017), p. 0067.

[44] Samuel J Gershman. "Do learning rates adapt to the distribution of rewards?" In: *Psychonomic bulletin & review* 22 (2015), pp. 1320–1327.

[45] Zahra Barakchian, Abdol-Hossein Vahabie, and Majid Nili Ahmadabadi. "Implicit counterfactual effect in partial feedback reinforcement learning: behavioral and modeling approach". In: *Frontiers in Neuroscience* 16 (2022), p. 631347.

[46] JOHN P O'DOHERTY, Alan Hampton, and Hackjin Kim. "Model-based fMRI and its application to reward learning and decision making". In: *Annals of the New York Academy of sciences* 1104.1 (2007), pp. 35–53.

[47] Wolfram Schultz, Peter Dayan, and P. Read Montague. "A Neural Substrate of Prediction and Reward". In: *Science* 275.5306 (Mar. 1997), pp. 1593–1599. ISSN: 1095-9203. DOI: 10.1126/science.275.5306.1593. URL: http://dx.doi.org/10.1126/science.275.5306.1593.

[48] Paul W Glimcher. "Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis". In: *Proceedings of the National Academy of Sciences* 108.supplement_3 (2011), pp. 15647–15654.

[49] John P O'doherty. "Reward representations and reward-related learning in the human brain: insights from neuroimaging". In: *Current opinion in neurobiology* 14.6 (2004), pp. 769–776.

[50] Yael Niv. "Reinforcement learning in the brain". In: *Journal of Mathematical Psychology* 53.3 (2009), pp. 139–154.

[51] Nathaniel D Daw et al. "Trial-by-trial data analysis using computational models". In: *Decision making, affect, and learning: Attention and performance XXIII* 23.1 (2011), pp. 3–38.

[52] John P O'Doherty et al. "Temporal difference models and reward-related learning in the human brain". In: *Neuron* 38.2 (2003), pp. 329–337.

[53] Henry W Chase et al. "Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis". In: *Cognitive, affective, & behavioral neuroscience* 15 (2015), pp. 435–459.

[54] Robin Paul Wilson et al. "The neural substrate of reward anticipation in health: a meta-analysis of fMRI findings in the monetary incentive delay task". In: *Neuropsychology review* 28.4 (2018), pp. 496–506.

[55] Thorsten Kahnt et al. "Disentangling neural representations of value and salience in the human brain". In: *Proceedings of the National Academy of Sciences* 111.13 (2014), pp. 5000–5005.

[56] MFS Rushworth et al. "Action sets and decisions in the medial frontal cortex". In: *Trends in cognitive sciences* 8.9 (2004), pp. 410–417.

[57] Uta Noppeney, Dirk Ostwald, and Sebastian Werner. "Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex". In: *Journal of Neuroscience* 30.21 (2010), pp. 7434–7446.

[58] Mohamed L Seghier. "The angular gyrus: multiple functions and multiple subdivisions". In: *The Neuroscientist* 19.1 (2013), pp. 43–61.

[59] Andrea E Cavanna and Michael R Trimble. "The precuneus: a review of its functional anatomy and behavioural correlates". In: *Brain* 129.3 (2006), pp. 564–583.

[60] Atsushi Yamaguchi and Tatsuya Jitsuishi. "Structural connectivity of the precuneus and its relation to resting-state networks". In: *Neuroscience Research* 209 (2024), pp. 9–17.

[61] Tatsuya Jitsuishi and Atsushi Yamaguchi. "Characteristic cortico-cortical connection profile of human precuneus revealed by probabilistic tractography". In: *Scientific reports* 13.1 (2023), p. 1936.

[62] R.S. Sutton and A.G. Barto. "Reinforcement Learning: An Introduction". In: *IEEE Transactions on Neural Networks* 9.5 (1998), pp. 1054–1054. DOI: 10.1109/TNN.1998.712192.

[63] Oscar Bartra, Joseph T McGuire, and Joseph W Kable. "The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value". In: *Neuroimage* 76 (2013), pp. 412–427.

[64] Florent Meyniel, Maxime Maheu, and Stanislas Dehaene. "Human inferences about sequences: A minimal transition probability model". In: *PLoS computational biology* 12.12 (2016), e1005260.

[65] John M Pearce and Geoffrey Hall. "A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli." In: *Psychological review* 87.6 (1980), p. 532.

[66] William W Seeley. "The salience network: a neural system for perceiving and responding to homeostatic demands". In: *Journal of Neuroscience* 39.50 (2019), pp. 9878–9882.

[67] Scott Marek and Nico UF Dosenbach. "The frontoparietal network: function, electrophysiology, and importance of individual precision mapping". In: *Dialogues in clinical neuroscience* 20.2 (2018), pp. 133–140.

[68] Russell A Poldrack. "Precision Neuroscience: Dense Sampling of Individual Brains". In: 95.4 (2017). ISSN: 0896-6273. DOI: 10.1016/j.neuron.2017.08.002.

[69] Zach Ladwig et al. "Precision fMRI reveals densely interdigitated network patches with conserved motifs in the lateral prefrontal cortex". In: *bioRxiv* (2025), p. 2025.07.24.666468. DOI: 10.1101/2025.07.24.666468.

[70] Yu Song et al. "Functional hierarchy of the angular gyrus and its underlying genetic architecture". In: *Human Brain Mapping* 44.7 (2023), pp. 2815–2828.

[71] Kathleen S Rockland. "Angular gyrus: an anatomical case study for association cortex". In: *Brain Structure and Function* 228.1 (2023), pp. 131–143.

[72] Karl Friston. "A theory of cortical responses". In: *Philosophical transactions of the Royal Society B: Biological sciences* 360.1456 (2005), pp. 815–836.

[73] Jeffrey R Binder and Rutvik H Desai. "The neurobiology of semantic memory". In: *Trends in cognitive sciences* 15.11 (2011), pp. 527–536.

[74] Jeffrey R Binder et al. "Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies". In: *Cerebral cortex* 19.12 (2009), pp. 2767–2796.

[75] Sarah V Di Pietro et al. "Disentangling influences of dyslexia, development, and reading experience on effective brain connectivity in children". In: *NeuroImage* 268 (2023), p. 119869.

[76] Barry Horwitz, Judith M Rumsey, and Brian C Donohue. "Functional connectivity of the angular gyrus in normal reading and dyslexia". In: *Proceedings of the National Academy of Sciences* 95.15 (1998), pp. 8939–8944.

[77] Pedro M Paz-Alonso et al. "Neural correlates of phonological, orthographic and semantic reading processing in dyslexia". In: *NeuroImage: Clinical* 20 (2018), pp. 433–447.

[78] Karen Gross-glenn et al. "Positron emission tomographic studies during serial word-reading by normal and dyslexic adults". In: *Journal of clinical and experimental neuropsychology* 13.4 (1991), pp. 531–544.

[79] Fabio Richlan, Martin Kronbichler, and Heinz Wimmer. "Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies". In: *Human brain mapping* 30.10 (2009), pp. 3299–3308.

[80] Fabio Richlan, Martin Kronbichler, and Heinz Wimmer. "Meta-analyzing brain dysfunctions in dyslexic children and adults". In: *Neuroimage* 56.3 (2011), pp. 1735–1742.

[81] Fahd Yazin et al. "Fragmentation and multithreading of experience in the default-mode network". In: *Nature Communications* 16.1 (2025), p. 8401.

[82] Steven G Dickstein et al. "The neural correlates of attention deficit hyperactivity disorder: An ALE meta-analysis". In: *Journal of child psychology and psychiatry* 47.10 (2006), pp. 1051–1062.

[83] Samuele Cortese et al. "Toward systems neuroscience of ADHD: a meta-analysis of 55 fMRI studies". In: *American journal of psychiatry* 169.10 (2012), pp. 1038–1055.

[84] Kaitlyn E May and Rajesh K Kana. "Frontoparietal network in executive functioning in autism spectrum disorder". In: *Autism Research* 13.10 (2020), pp. 1762–1777.

[85] José M Maisog et al. "A meta-analysis of functional neuroimaging studies of dyslexia". In: *Annals of the new York Academy of Sciences* 1145.1 (2008), pp. 237–259.

[86] Aoife Lonergan et al. "A meta-analysis of executive functioning in dyslexia with consideration of the impact of comorbid ADHD". In: *Journal of Cognitive Psychology* 31.7 (2019), pp. 725–749.

[87] Courtney Pollack, Gigi Luk, and Joanna A Christodoulou. "A meta-analysis of functional reading systems in typically developing and struggling readers across different alphabetic languages". In: *Frontiers in Psychology* 6 (2015), p. 191.

[88] Fabio Richlan. "The functional neuroanatomy of letter-speech sound integration and its relation to brain abnormalities in developmental dyslexia". In: *Frontiers in human neuroscience* 13 (2019), p. 21.

[89] Anastasia Christakou et al. "Disorder-specific functional abnormalities during sustained attention in youth with attention deficit hyperactivity disorder (ADHD) and with autism". In: *Molecular psychiatry* 18.2 (2013), pp. 236–244.

[90] Anna Martin, Martin Kronbichler, and Fabio Richlan. "Dyslexic brain activation abnormalities in deep and shallow orthographies: A meta-analysis of 28 functional neuroimaging studies". In: *Human brain mapping* 37.7 (2016), pp. 2676–2699.

[91] Vincent Walsh and Alvaro Pascual-Leone. *Transcranial magnetic stimulation: a neurochronometrics of mind*. MIT press, 2003.

[92] Alvaro Pascual-Leone, Vincent Walsh, and John Rothwell. "Transcranial magnetic stimulation in cognitive neuroscience–virtual lesion, chronometry, and functional connectivity". In: *Current opinion in neurobiology* 10.2 (2000), pp. 232–237.

[93] Tal Burt, Sarah H Lisanby, and Harold A Sackeim. "Neuropsychiatric applications of transcranial magnetic stimulation: a meta analysis". In: *International Journal of Neuropsychopharmacology* 5.1 (2002), pp. 73–103.

[94] Chris Rorden and Hans-Otto Karnath. "Using human brain lesions to infer function: a relic from a past era in the fMRI age?" In: *Nature Reviews Neuroscience* 5.10 (2004), pp. 812–819.

[95] Nir Grossman et al. "Noninvasive deep brain stimulation via temporally interfering electric fields". In: *cell* 169.6 (2017), pp. 1029–1041.

[96] Davide Folloni et al. "Manipulation of subcortical and deep cortical activity in the primate brain using transcranial focused ultrasound stimulation". In: *Neuron* 101.6 (2019), pp. 1109–1116.

[97] Yechiel Levkovitz et al. "Efficacy and safety of deep transcranial magnetic stimulation for major depression: a prospective multicenter randomized controlled trial". In: *World Psychiatry* 14.1 (2015), pp. 64–73.

[98] Nina Raduner et al. "Growing minds, integrating senses: Neural and computational insights into age-related changes in audio-visual and tactile-visual learning in children". In: *Developmental Cognitive Neuroscience* 76 (2025), p. 101622. ISSN: 1878-9293. DOI: https://doi.org/10.1016/j.dcn.2025.101622. URL: https://www.sciencedirect.com/science/article/pii/S1878929325001185.

[99] Oscar Esteban et al. "fMRIPrep: a robust preprocessing pipeline for functional MRI". In: *Nature methods* 16.1 (2019), pp. 111–116.

[100] Krzysztof Gorgolewski et al. "Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python". In: *Frontiers in neuroinformatics* 5 (2011), p. 13.