Cognitive Characteristics and Neural Mechanisms of Multichannel Category Learning: EEG and DDM Evidence

Authors: Wu Jie, Che Zixuan, Wu Jie

Date: 2025-07-04T16:21:06+00:00

Abstract

The cognitive characteristics and neural mechanisms of multi-channel category learning are of critical importance for elucidating the principles of cross-modal knowledge representation. This study combines event-related potential (ERP) technology with the drift-diffusion model to systematically investigate the cognitive characteristics and neural mechanisms of multi-channel category learning. Behavioral results showed that, compared to the early learning stage, the middle and late learning stages exhibited significant increases in accuracy and drift rate, significant decreases in response time, and a shift in the decision starting point toward the correct option. At the neural level, the middle and late learning stages elicited changes in the amplitudes of N1, P1, N250, FSP (Frontal Selection Positivity), and LPC (Late Positive Component); time-frequency analysis revealed significant attenuation in Theta, Alpha, and Delta frequency bands. Regression analysis indicated that N250-FSP amplitude and Theta oscillations jointly explained variance in drift rate, while P1, N250-FSP, and LPC could predict shifts in the decision starting point. The study demonstrates that learning training optimizes decision efficacy through dual mechanisms: (1) enhanced information accumulation rate is associated with reduced N250-FSP amplitude and attenuated Theta band power; (2) shifts in the decision starting point are driven by the synergistic interaction of early perceptual encoding (P1), feature discrimination (N250-FSP), and memory retrieval (LPC).

Full Text

The Cognitive Characteristics and Neural Mechanisms of Multisensory Category Learning: EEG and Drift-Diffusion Model Evidence

WU Jie, CHE Zixuan

School of Psychology, Fujian Normal University, Fuzhou 350117, China

Abstract

Multisensory category learning represents a fundamental cognitive function that is crucial for understanding the principles of cross-modal knowledge representation. This study systematically investigated the cognitive characteristics and neural mechanisms underlying multisensory category learning by integrating event-related potential (ERP) techniques with drift-diffusion modeling. Behavioral results demonstrated that compared with the early learning stage, both middle and later stages exhibited significantly improved accuracy and drift rates, reduced reaction times, and a decision starting point biased toward correct options. At the neural level, middle and later learning stages elicited amplitude changes in N1, P1, N250, FSP (Frontal Selection Positivity), and LPC (Late Positive Component) components. Time-frequency analysis revealed significant power attenuation in the Theta, Alpha, and Delta frequency bands. Regression analyses indicated that N250-FSP amplitude and Theta oscillations jointly explained variance in drift rates, while P1, N250-FSP, and LPC predicted shifts in decision starting points. These findings suggest that learning training optimizes decision efficiency through dual mechanisms: (1) enhanced information accumulation rates correlate with reduced N250-FSP amplitudes and Theta band power attenuation; (2) decision starting point shifts are driven by the synergistic interaction of early perceptual encoding (P1), feature discrimination (N250-FSP), and memory retrieval (LPC).

Keywords: Multisensory, Category learning, Drift-Diffusion Model, EEG

Introduction

Category learning in multisensory environments constitutes a foundational human cognitive function that integrates cross-modal information to classify objects and form stable memory representations (Li et al., 2012). This cognitive mechanism extracts common features across multisensory inputs, transforming concrete experiences into abstract conceptual frameworks that significantly optimize cognitive resource allocation efficiency. Leveraging these advantages, individuals can transfer existing category knowledge to novel objects and environments, thereby enhancing task performance in new situations and improving environmental adaptability (Seger & Miller, 2010). Empirical research demonstrates that category learning plays a pivotal role in higher-order cognitive activities including object recognition, logical reasoning, and complex problem-

solving. In clinical medicine, precise disease classification systems form the basis of diagnostic and treatment decisions, while in ecological conservation, scientific waste classification mechanisms represent crucial practices for sustainable development. Therefore, systematically revealing the acquisition mechanisms and formation principles of multisensory category knowledge holds not only theoretical value for understanding the essence of human cognitive processing but also significant practical implications for enhancing individual survival capabilities and optimizing social decision-making practices.

Multisensory category learning differs fundamentally from unisensory category learning. The former requires not only independent processing of modality-specific representations (e.g., visual, auditory, and olfactory information) but also hierarchical integration of cross-modal information to ultimately achieve categorical decisions through intermodal synthesis. Notably, its cognitive mechanisms diverge from basic processing stages such as multisensory information detection, memory, and associative learning, transcending memory encoding at the level of single features or exemplars. The core challenge lies in extracting common features across multisensory objects and elevating concrete experiences to abstract conceptual representations (Li et al., 2012). This cognitive-level transition implies that principles derived from unisensory learning paradigms and memory encoding mechanisms revealed in multisensory processing studies cannot be directly extrapolated to multisensory category learning (Seger & Miller, 2010). This research domain requires an independent theoretical framework to uncover its unique neurocognitive mechanisms.

Unisensory Category Learning

Neuroimaging studies of visual category learning have revealed that classification processing primarily involves two stages: perceptual encoding and categorical feature discrimination (Freedman et al., 2003; Jiang et al., 2007; Jiang et al., 2018; Scholl et al., 2014). The N1 component, emerging as the first negative deflection 150–200 ms post-stimulus, is significantly associated with stimulus detection (Busse et al., 2005) and feature encoding (Sinnett et al., 2007), providing direct evidence for early shape and feature perception in category learning (Scholl et al., 2014). Freedman et al. (2003) employed intracranial recordings in macaque categorization tasks, finding that the inferior temporal cortex exhibited stimulus-specific activation within 100 ms for shape features, while the prefrontal cortex selectively processed abstract categorical dimension information within 200 ms—temporal dynamics that corroborate hierarchical model assumptions.

During advanced cognitive processing stages, the N250 and FSP (Frontal Selection Positivity) components are closely linked to semantic concept processing. Experimental evidence demonstrates that category-relevant stimuli elicit stronger N250 and FSP components at 250 ms compared with irrelevant stimuli (Folstein, Monfared, et al., 2017). Scholl et al. (2014) further observed that between-category stimuli evoked significantly more negative potentials in

the parietal cortex than within-category stimuli during the 200–300 ms time window. These electrophysiological features collectively indicate that category judgment follows a hierarchical processing pattern from primary feature perception to high-level categorical decision-making.

Previous research evidence indicates that unisensory category learning comprises early feature perception and late categorical decision stages (Freedman et al., 2003; Jiang et al., 2007; Jiang et al., 2018; Scholl et al., 2014). Unlike unisensory category learning, multisensory category learning additionally involves a critical stage of perceiving and integrating information across different modalities. However, which specific ERP components correlate with this processing stage remains unclear.

Multisensory Information Integration, Memory, and Paired Association Learning

ERP studies of multisensory information processing have revealed that, at the detection level, multisensory stimuli elicit significantly enhanced early evoked potentials within 100 ms compared with unisensory stimuli (Cappe et al., 2012; Ghazanfar & Schroeder, 2006; Giard & Peronnet, 1999; Senkowski et al., 2011; Van der Burg et al., 2011), with semantically congruent information producing larger amplitudes in the 180–210 ms time window (Hu et al., 2012). Memory research demonstrates that multisensory encoding enhances recognition accuracy, with neural manifestations of amplitude enhancement in 100–130 ms and 270–316 ms time windows (Thelen et al., 2012; Thelen et al., 2014). In cross-modal paired association learning, visuo-tactile associative learning modulates N400 and Late Posterior Negativity (LPN) amplitudes, with significant differences between N400 and LPN components in the well-learned versus initial learning stages (Gui et al., 2017).

Neural oscillation studies of multisensory information processing have shown that, at the detection level, verbal cues preceding odor delivery induce phase locking between low-frequency (<30 Hz) oscillations in auditory and olfactory cortices (Zhou et al., 2019). At the multisensory attention level, research examining attention across visual, auditory, and audiovisual conditions found that Theta band (4–7 Hz) activity is associated with multisensory attention, while Alpha band (8–13 Hz) activity correlates with auditory attention (Keller et al., 2017). Another study employing an audiovisual semantic discrimination task requiring attention to both modalities demonstrated that Theta band activity contributes to early attention allocation regulation, whereas Delta band activity facilitates later attention allocation in audiovisual integration (Yang et al., 2024). In multisensory memory research, manipulating Theta phase synchronization across visual and auditory stimuli revealed that synchronous conditions produced significantly superior multisensory associative memory performance compared with asynchronous conditions (90°, 180°, 270° phase differences). This finding demonstrates that Theta oscillations enhance memory integration by coordinating temporal windows of cross-modal neural activity

(Clouter et al., 2017).

Multisensory research demonstrates that information detection primarily correlates with early components (N1/P1), while higher-order cognitive processes such as memory and learning involve coordinated activity of late components (e.g., N400 and LPN). Furthermore, multisensory attention and memory are associated with Alpha, Theta, and Delta frequency bands. However, multisensory information detection, memory, and paired association learning have focused on detecting, memorizing, or jointly processing specific individual stimuli, with less emphasis on abstract information summarization such as rule extraction. In contrast, multisensory category learning represents a more complex cognitive process that encompasses not only perception and memory of individual stimuli but also integration of abstract rules and concepts. Therefore, investigating the cognitive characteristics and neural mechanisms of multisensory category learning holds significant theoretical importance.

Drift-Diffusion Model and Multisensory Category Learning

Behavioral studies of multisensory category learning indicate that learning progression significantly improves task accuracy and shortens reaction times (Wu et al., 2021). Although this pattern resembles most learning tasks, it cannot reveal which psychological processes underlie accuracy improvements and RT reductions—whether from shortened non-decision time, enhanced information integration rates, more liberal decision thresholds, or initial bias toward correct options at the decision outset. The Drift-Diffusion Model (DDM) deconstructs decision processes into four core parameters: drift rate (evidence accumulation rate), decision boundary (response caution), non-decision time (stimulus encoding duration), and starting point (response bias) (Yuan et al., 2023; Mąka et al., 2023). These parameters reflect psychological processes in category learning: drift rate correlates with behavioral performance and represents the rate of integrating visual and auditory features; decision boundary reflects conservatism in categorical judgments; non-decision time maps onto early perceptual processing efficiency or motor response time; and starting point manifests prior experience modulation of category judgments. Empirical research shows that for difficult tasks, younger adults exhibit superior evidence accumulation efficiency and decision speed compared with older adults in multisensory tasks; for easy tasks, multisensory advantages in evidence accumulation increase with age (Bolam et al., 2024). Cross-modal (visual vs. auditory) category learning research found slower learning rates for visual than auditory learning in early stages (Roark et al., 2021), with age-related cognitive strategy differentiation—adults' overall advantage over children attributed to enhanced information processing capacity, while their superior performance in explicit visual and implicit auditory category learning related to less cautious correct responses (Roark et al., 2023).

However, existing research has not clarified the dynamic associations between category knowledge acquisition and specific cognitive parameters, necessitating systematic investigation of plastic changes in decision mechanisms during

learning progression. At the neural mechanism level, although studies have confirmed that N1 (shape and feature perception), N250/FSP (categorical feature discrimination), N400/LPN (cross-modal paired learning), and Alpha, Theta, and Delta bands (cross-modal information processing) show specific representations in unisensory category learning and multisensory processing, how different ERP components and band power influence information accumulation rates and decision caution during multisensory category learning remains unknown.

This study integrates high-temporal-resolution EEG technology with cognitive modeling (DDM) to systematically parse the cognitive characteristics and neural mechanisms of multisensory category learning by constructing dynamic association models between ERP components/band power and DDM parameters from dual dimensions of neural oscillation and computational cognitive modeling. Based on evidence that unisensory category learning involves shape perception and categorical decision stages, and that multisensory information detection, memory, and learning involve multiple ERP components, we propose a hierarchical neural computational model hypothesis: multisensory category learning involves three stages—shape and feature perception, categorical feature integration, and categorical decision-making—with early ERP components associated with shape/feature perception and late components linked to categorical feature integration and decision-making. Additionally, considering that Theta, Alpha, and Delta bands are closely related to attention and memory processes, we hypothesize that power in these bands changes with learning progression, and that changes in ERP amplitudes and band power can predict dynamic variations in core DDM parameters such as drift rate.

2.1 Participants

Sample size was determined a priori using G*Power 3.1 software (Faul et al., 2007). Based on a repeated-measures design (effect size f = 0.25, $\alpha = 0.05$, power = 0.80), the minimum required sample size was 21 participants. Following comparable research designs (Folstein, Fuller, et al., 2017; Folstein, Monfared, et al., 2017; Liu et al., 2024), we recruited 30 healthy university students (7 males, mean age 20.6 $\pm$ 1.32 years). To ensure data validity and accurate cognitive characterization, we excluded 6 participants who failed to achieve $ $80% accuracy in the final experimental block (Wu et al., 2021; Liu et al., 2021), retaining 24 valid datasets. All participants reported normal or corrected-to-normal vision and normal hearing, with no prior experience in similar experiments. The experimental protocol was approved by the Ethics Committee of the School of Psychology, Fujian Normal University (Approval No.: PSY240073). Participants provided written informed consent prior to the experiment and received ¥50 compensation upon completion.

Experimental Design

This study employed a single-factor within-subjects design with learning stage as the independent variable, comprising three conditions: early, middle, and late learning stages. Dependent variables systematically included behavioral and neurophysiological metrics: at the behavioral level, we collected accuracy and reaction time data and extracted four core cognitive parameters from the drift-diffusion model (drift rate, non-decision time, starting point position, decision boundary); at the neurophysiological level, we simultaneously recorded event-related potential (ERP) component characteristics and EEG band power changes.

2.3 Experimental Materials

Visual and auditory stimuli were selected from the multisensory material library constructed by Wu et al. (2021). Visual stimuli generation followed the dual-prototype morphing paradigm proposed by Folstein et al. (2013), using two vehicle images as prototype stimuli. Through GTK morph software, we generated continuous visual stimulus sequences by morphing the prototype images. As shown in [Figure 1: see original paper], the two prototype images were divided into N corresponding grid units, and morphing parameters were controlled to generate continuous stimulus sequences from Prototype A (100% morph) to Prototype B (0% morph). The final visual material library comprised 99 morphed images with morph levels ranging from 1% to 99% in 1% increments. All images maintained a resolution of 600$\times$400 pixels, corresponding to horizontal and vertical visual angles of 9.9° and 9.1°, respectively, ensuring image clarity and precise detail presentation.

Auditory stimuli were synthesized using the STRAIGHT speech synthesis toolbox (Kawahara & Matsui, 2003), constructing continuous acoustic spaces based on two vehicle horn prototypes selected from a standard sound effects library (http://sc.chinaz.com/yinxiao/). After extracting fundamental frequency parameters (F01 and F02) via the STRAIGHT system, we employed linear interpolation algorithms to generate 99 auditory stimuli with morph gradients synchronized to visual materials (1%–99%). All auditory stimuli underwent sound pressure level normalization using MP3Gain, with final presentation intensity uniformly set at 50 dB SPL using binaural balanced output. This processing protocol effectively controlled physical feature variations in auditory stimuli, ensuring perceptual comparability.

The experiment constructed a four-category task through visuo-auditory dual-channel integration. As illustrated in [Figure 1: see original paper], we established a two-dimensional classification coordinate system based on visual and auditory stimulus dimensions, using median values (50%) in both dimensions as the origin to divide four decision quadrants, each corresponding to an independent category. This generated four categories through full factorial combination, producing 2,304 total stimuli. Category generation rules were: (1)

Category A combined Visual Prototype A (51%–99% morph) with Auditory Prototype A (51%–99% morph); (2) Category B combined Visual Prototype B (1%–49% morph) with Auditory Prototype A (51%–99% morph); (3) Category C combined Visual Prototype B (1%–49% morph) with Auditory Prototype B (1%–49% morph); (4) Category D combined Visual Prototype A (51%–99% morph) with Auditory Prototype B (1%–49% morph). The experiment employed a blocked design, with each block containing 40 trials randomly selected from the stimulus library. Each participant completed 9 blocks (360 trials total), with all stimuli presented equally across blocks.

[Figure 1: see original paper]. Example stimuli for the four categories in multisensory category learning.

Experimental Procedure

The trial structure is illustrated in [Figure 2: see original paper]. A single trial comprised: (1) Fixation phase: a central cross presented for 700–900 ms (random duration); (2) Stimulus presentation: 600 ms bimodal stimulus (visual + auditory compound); (3) Response phase: following stimulus offset, a 400 ms blank screen preceded a free-response period during which participants classified stimuli via four-key responses (Categories A–D mapped to D/F/J/K keys, counterbalanced across participants); (4) Feedback phase: 1000 ms correctness feedback after keypress, followed by a 500 ms intertrial interval. Instructions explicitly informed participants to accumulate categorical rules across trials, with initial random guessing acceptable but feedback utilization required to gradually form stable category representations.

[Figure 2: see original paper]. Experimental procedure for the learning process.

2.5 EEG Data Collection and Analysis

EEG signals were recorded using a NeuroScan 64-channel system with electrode placement following the international 10–20 system. Data were synchronized via SCAN software, referenced to the left mastoid (M1) with AFz as ground. All electrode impedances were maintained below 5 kΩ. Vertical and horizontal electrooculograms were simultaneously recorded for artifact detection, with signals amplified and digitized at 1,000 Hz sampling rate.

Data preprocessing was conducted using EEGLAB (Delorme & Makeig, 2004) and FieldTrip (Oostenveld et al., 2011) toolboxes: (1) Re-referencing: converted reference to averaged bilateral mastoids (M1/M2); (2) Filtering: 0.1–120 Hz bandpass filter removed high-frequency noise (>120 Hz) and low-frequency drift (<0.1 Hz), with a 50 Hz notch filter (49–51 Hz) eliminating power-line interference; (3) Epoching: data segments from –200 ms to 1,000 ms relative to stimulus onset were extracted, with baseline correction using the –200 to 0 ms prestimulus interval; (4) Artifact correction: independent component analysis (ICA) removed ocular and blink-related components.

Statistical analysis employed FieldTrip for condition-wise averaging. To investigate neural mechanism changes induced by multisensory category learning, we contrasted EEG signals across early, middle, and late learning stages. Cluster-based permutation tests (Maris & Oostenveld, 2007) were applied, conducting paired two-tailed t-tests on amplitudes and band power across all electrode-time points within a 1,000 ms post-stimulus window. A null distribution was generated via 5,000 permutations, with αthresh = 0.025 as the cluster-forming threshold to control for multiple comparisons and identify statistically significant spatiotemporal clusters.

2.6 Model Fitting

We employed the Python-based Hierarchical Drift-Diffusion Model toolbox HDDM (Wiecki et al., 2013) to construct hierarchical Bayesian DDMs analyzing behavioral differences across early, middle, and late learning stages. Core parameters included drift rate (v), decision boundary (a), non-decision time (t), and starting point bias (z), which were allowed to vary across experimental conditions. To improve model fit, we also estimated drift rate variability (sv), non-decision time variability (st), and starting point variability (sz) as free parameters, with 5% extreme RT data trimmed to control outlier interference.

Parameter estimation employed Markov Chain Monte Carlo (MCMC) Bayesian inference, executing 5,000 iterations with the first 500 as burn-in to ensure chain convergence. Convergence was assessed via visual inspection of trace plots, posterior distribution histograms, and autocorrelation functions. Model fit was compared using Deviance Information Criterion (DIC), selecting the optimal model with minimal DIC value, which offers statistical advantages in balancing model complexity and fit.

Based on the optimal model, we extracted posterior distributions of four key parameters for each participant across three learning stages. Condition differences were tested using Bayesian inference on posterior distributions, directly computing posterior probability differences to avoid limitations of frequentist statistics. All analyses were implemented via HDDM's built-in Bayesian computation methods, ensuring robust and reliable parameter estimation.

3.1 Behavioral Results

To examine performance changes across learning stages, we conducted repeated-measures ANOVA on accuracy and reaction time for early (first 3 blocks, 120 trials), middle (middle 3 blocks, 120 trials), and late (final 3 blocks, 120 trials) learning stages, as shown in [Figure 3: see original paper]. Statistical analysis revealed significant main effects for accuracy, $F(2, 46) = 132.30$, $p < 0.001$, $^2 = 0.85$, and reaction time, $F(2, 46) = 7.58$, $p = 0.001$, $^2 = 0.25$. Post-hoc simple effects analysis indicated that middle and late stage accuracy was significantly higher than early stage accuracy ($ps < 0.001$), with late stage accuracy significantly exceeding middle stage accuracy ($p < 0.001$). Reaction

times in middle (p = 0.001) and early (p = 0.05) stages were significantly longer than in the late stage. These results demonstrate that systematic learning enabled participants to successfully establish effective discrimination among four categories.

To further investigate the psychological mechanisms underlying accuracy improvements, we examined whether these effects stemmed from decision boundary adjustments (changes in decision caution), enhanced information accumulation rates, or reduced non-decision time (stimulus encoding duration). Given that traditional behavioral metrics (accuracy/RT) cannot effectively distinguish these potential mechanisms, we employed drift-diffusion modeling to quantify dynamic changes in cognitive parameters during learning.

[Figure 3: see original paper]. Accuracy and reaction time across early, middle, and late learning stages.

3.2 HDDM Model Analysis

We constructed seven computational models with different parameter combinations, with the full-parameter model including v (drift rate), a (decision boundary), t (non-decision time), z (starting point), and sv, st, sz variability parameters demonstrating optimal fit. Using this optimal model architecture, we fitted behavioral data (accuracy and RT) across early, middle, and late conditions, successfully estimating individual participant values for four core parameters (v, a, t, z) under each condition. Group difference tests based on direct posterior distribution estimation ([Figure 4: see original paper]) yielded key findings:

First, regarding information processing efficiency, drift rates in the early stage were significantly lower than in middle (p < 0.001) and late (p < 0.001) stages, with middle stage drift rates also significantly lower than late stage (p < 0.001). This indicates that as learning progressed, individuals integrated more effective information per unit time, achieving significant enhancement in information accumulation rates. Second, regarding decision starting point position, early stage differed significantly from middle (p = 0.03) and late (p = 0.008) stages, with middle and late stages showing significant bias toward correct options. This finding mechanistically corresponds to the behaviorally observed accuracy improvements, revealing systematic modulation of decision bias by learning progression. Notably, no significant differences emerged across conditions in decision boundary or non-decision time, suggesting that learning-induced changes primarily affected information processing efficiency and decision bias rather than decision caution or non-decision time.

[Figure 4: see original paper]. Probability density distributions of four DDM parameters in multisensory category learning. Panel (A) represents drift rate; (B) non-decision time; (C) decision boundary; (D) starting point position.

3.3.1 Time-Domain Results

We used permutation tests to contrast EEG amplitude differences across early, middle, and late learning stages ([Figure 5: see original paper]). In early processing, significant differences between early and late stages emerged in 100–180 ms (N1 and P1, p = 0.01) across electrodes CZ, C2, CP3, CP1, CPZ, CP2, CP4, CP6, TP8, P7, P5, P3, P1, PZ, P2, P4, P6, P8, PO7, PO5, PO3, POZ, PO4, PO6, PO8, O1, OZ, O2. Early versus middle stages differed significantly in 100–167 ms (N1 and P1, p = 0.03) across electrodes C5, C3, C1, CZ, C2, CP5, CP3, CP1, CPZ, CP2, CP4, CP6, TP8, P7, P5, P3, P1, PZ, P2, P4, P6, P8, PO7, PO5, PO3, POZ, PO4, PO6, PO8, O1, OZ, O2. Middle versus late stages differed in 128–172 ms (N1 and P1, p = 0.05) across electrodes CP6, P1, PZ, P2, P4, P6, PO7, PO5, PO3, POZ, PO4, PO6, PO8, O1, OZ, O2.

In late processing, early versus late stages differed significantly in 210–326 ms (FSP, p = 0.04) across electrodes TP8, P6, P8, PO4, PO8, OZ, O2; 240–492 ms (N250, p < 0.001) across frontal, frontocentral, central, and parietal electrodes; and 493–1000 ms (LPC, p < 0.001) across central, centroparietal, and parietal electrodes. Early versus middle stages showed marginal significance in 231–328 ms (FSP, p = 0.056) and significant differences in 198–495 ms (N250, p < 0.001) and 496–1000 ms (LPC, p < 0.001).

[Figure 5: see original paper]. ERP waveforms and topographic difference maps across learning stages. (A) Mean amplitudes at CZ, CPZ, PZ, POZ, and OZ; (B) Mean amplitudes at FZ, FCZ, CZ, CPZ, PZ, and POZ; (C) Mean amplitudes at TP8, P8, PO8, and O2; (D) Mean amplitudes at CZ, CPZ, PZ, and POZ.

3.3.2 Frequency-Domain Results

Time-frequency analysis revealed significant power differences in the 3–13 Hz range between early versus middle and late learning stages ([Figure 6: see original paper]). Specifically, compared with the early stage, the late stage exhibited significant power attenuation in Delta (3 Hz), Theta (4–7 Hz), and Alpha (8–13 Hz) bands (p = 0.002), with this power reduction distributed globally across electrodes. Similarly, the middle stage showed significant power attenuation in Delta, Theta, and Alpha bands (p = 0.009) compared with the early stage, also with widespread topographical distribution.

[Figure 6: see original paper]. Band power changes across learning stages: (A) Late minus early stage power differences; (B) Middle minus early stage power differences.

3.4 EEG Signals Predicting Drift Rate and Starting Point Bias

ERP Signals Predicting Drift Rate and Starting Point Bias

Integrating time-domain analysis with computational modeling, we systematically examined neurocomputational associations between ERP components and decision dynamics parameters. Time-domain analysis revealed significant differences across learning stages in early perceptual components (P1, N1), categorical feature processing components (N250, FSP), and memory retrieval components (LPC). DDM analysis further demonstrated significant dissociations across stages in information accumulation rate (drift rate) and response bias (starting point) parameters.

Drawing on previous theoretical frameworks (Cavanagh et al., 2011; Herz et al., 2016) and integrating time-domain and DDM results, we constructed multiple regression models to examine ERP prediction of decision parameters: Drift Rate/Starting Point $= \alpha + eP1{\times}P1 + eN1{\times}N1 + eN250{\times}N250 + eFSP{\times}FSP + eLPC{\times}LPC$. Regression weights eP1, eN1, eN250, eFSP, and eLPC tested the effects of P1, N1, N250, FSP, and LPC on drift rate and starting point, with α as intercept. Analyses of component means and peaks revealed that N250 (peak: $p = 0.001$; mean: $p < 0.001$) and FSP (peak: $p < 0.001$; mean: $p < 0.001$) significantly negatively predicted drift rate, indicating that amplitude reductions in these components covaried with enhanced information accumulation rates. P1 (peak: $p = 0.06$; mean: $p = 0.07$), N1 (peak: $p = 0.07$; mean: $p = 0.93$), and LPC (peak: $p = 0.06$; mean: $p = 0.06$) showed non-significant predictive effects. In the starting point prediction model, P1 (peak: $p = 0.03$; mean: $p = 0.005$), N250 (peak: $p = 0.008$; mean: $p = 0.007$), FSP (peak: $p = 0.001$; mean: $p < 0.001$), and LPC (peak: $p = 0.03$; mean: $p = 0.016$) all reached significance, while N1 (peak: $p = 0.90$; mean: $p = 0.90$) was non-significant. These results suggest that early perceptual processing (P1), categorical feature discrimination (N250/FSP), and late memory retrieval (LPC) jointly shape response bias formation mechanisms. These findings indicate that multi-stage ERP components influence decision dynamics through hierarchical information processing: N250-FSP co-modulation may reflect regulation of information accumulation efficiency by categorical representation precision, while P1-LPC dynamic coupling represents the joint effect of perceptual encoding and memory retrieval on response bias.

[Figure 7: see original paper]. Effects of ERP components and band power on drift rate and starting point position. (A) ERP effects on drift rate; (B) Band power effects on drift rate; (C) ERP effects on starting point; (D) Band power effects on starting point.

Frequency Band Power Predicting Drift Rate and Starting Point Bias

Time-frequency analysis revealed significant differences in Delta (3 Hz), Theta (4–7 Hz), and Alpha (8–14 Hz) oscillations between early versus middle and late

stages. To examine each band's influence on DDM parameters, we constructed regression models following previous literature (Cavanagh et al., 2011): Drift Rate/Starting Point $= \alpha + e\tau \times \tau + e \times + e\alpha \times \alpha$, where $e\tau$, e, and $e\alpha$ tested Delta, Theta, and Alpha band effects. Statistical analysis revealed that Theta band activity significantly predicted drift rate ($p = 0.03$), while Delta ($p = 0.07$) and Alpha ($p = 0.76$) effects were non-significant. In the starting point bias regression model, no bands showed significant predictive effects (Theta: $p = 0.45$; Delta: $p = 0.52$; Alpha: $p = 0.17$). Regression weights and statistical significance are illustrated in [Figure 7: see original paper].

These results demonstrate that during multisensory category learning, participants showed significant progressive accuracy improvements and reduced reaction times in later stages. DDM-based cognitive modeling revealed that compared with early learning, middle and late stages exhibited significantly increased drift rates and starting point shifts toward correct options, with no condition differences in decision boundary or non-decision time. This suggests that learning-induced behavioral improvements primarily stemmed from enhanced evidence accumulation efficiency and adaptive adjustment of decision bias. Compared with early learning, middle and late stages elicited significant amplitude changes in N1, P1, N250, FSP, and LPC components, along with power reductions in Delta, Theta, and Alpha bands. Notably, N250-FSP complex and Theta band changes significantly predicted drift rate variance, while P1, N250-FSP, and LPC jointly predicted starting point shifts.

General Discussion

Hierarchical Processing Stages in Multisensory Category Learning

Synthesizing existing research (Jiang et al., 2007, 2018; Scholl et al., 2014), unisensory category learning follows a hierarchical processing pattern from simple to complex, primarily involving progressive transformation from perceptual encoding to categorical selection. In contrast, multisensory category learning involves more complex multimodal information integration mechanisms. Our ERP analysis revealed that multisensory learning progression induced significant changes in multiple characteristic EEG components (N1, P1, N250, FSP, LPC), with ERP components and band power significantly predicting core parameters such as drift rate. Based on multidimensional evidence, we propose that multisensory category learning may involve three progressive processing stages: initial perceptual encoding, feature refinement discrimination, and memory system activation. This multi-stage model reveals the hierarchical dynamic neural processing mechanisms in multisensory category learning.

Perceptual Encoding Stage Compared with early learning, middle and late stages induced significant N1 and P1 amplitude changes. As early perceptual processing markers (Cappe et al., 2012; Ghazanfar & Schroeder, 2006; Giard & Peronnet, 1999; Senkowski et al., 2011; Van der Burg et al., 2011), N1 and P1 likely reflect shape and feature perception of multisensory stimuli. Previous

visual category learning research found that high stimulus variability conditions elicited early amplitude changes compared with low variability conditions (Scholl et al., 2014), with N1 amplitude changes significantly associated with stimulus detection (Busse et al., 2005) and feature encoding (Sinnett et al., 2007). Other studies demonstrated that enhanced early P1 amplitude reflects refined encoding of multisensory features in primary sensory cortex (Mück et al., 2020). In this study, N1 and P1 amplitude changes may relate to altered shape and feature perception induced by multisensory category learning. Furthermore, multiple regression revealed that P1 significantly predicted correct-option bias shifts as learning progressed. These combined results demonstrate that early perceptual stages also play important roles in adaptive adjustments for correct responding. Notably, differences between early and later learning stages may reflect that participants initially employed separate information processing strategies, while post-learning enabled efficient classification through cross-modal integration. This transition theoretically aligns with the multisensory enhancement effect (Stevenson et al., 2014) and carries important theoretical implications for understanding multisensory information integration mechanisms.

Feature Discrimination Stage   The N250-FSP complex serves as a core neural marker for categorical feature-specific discrimination (Folstein, Monfared, et al., 2017). The amplitude reduction observed in middle and late learning stages provides crucial evidence for a feature discrimination stage in multisensory category learning progression. Previous research has established N250's specific association with complex object recognition (e.g., faces, animal categories) (Petrov, 2011; Tanaka et al., 2006). ERP studies further confirm that category-relevant stimuli elicit enhanced N250 and FSP amplitudes compared with irrelevant stimuli, reflecting refined processing of perceptual features (Folstein, Monfared, et al., 2017). Additionally, FSP demonstrates linear correlation with target feature learning level, particularly prominent in high-level classification tasks (Scott et al., 2006; 2008). Based on this evidence, we infer that N250-FSP may serve as an important index for multisensory categorical feature discrimination.

Moreover, the N250-FSP complex amplitude reduction, coupled with Delta, Alpha, and Theta band power attenuation, carries important theoretical significance. Reduced N250-FSP amplitude may indicate automation of categorical feature representations. Synchronized Delta, Alpha, and Theta power reductions suggest decreased attentional resource demands (Cavanagh & Frank, 2014; Keller et al., 2017), particularly Theta attenuation correlating with reduced memory load (Clouter et al., 2017). These power reductions collectively confirm optimized cognitive resource allocation. Our novel finding is that multisensory category learning induced N250/FSP amplitude suppression (rather than enhancement) and Theta band power attenuation, which systematically correlated with enhanced drift rates and correct-option bias. This inverse relationship indicates that as learning proficiency increased, individuals gradually formed automated processing patterns for cross-modal stimulus classification, manifested as reduced neural resource consumption during feature recognition

(N250/FSP amplitude reduction), decreased reliance on attentional monitoring and working memory maintenance during categorical decisions (Theta power reduction), and optimized information accumulation efficiency (drift rate enhancement). This finding provides new evidence for "perceptual learning compression theory" (Goldstone, 1998), suggesting that multisensory category training may enhance decision system processing efficiency by reducing cognitive load during feature discrimination.

Notably, this study reveals Theta oscillation's specific role in multisensory category learning. Theta band power significantly negatively predicted drift rate. This finding creates theoretical tension with Theta's classic functional roles in attentional resource allocation (Keller et al., 2017) and working memory maintenance (Itthipuripat et al., 2013; Rutishauser et al., 2010), where Theta enhancement rather than reduction typically predicts performance. Previous research shows that cross-modal attention tasks induce enhanced frontal Theta synchronization (Keller et al., 2017), and working memory load positively correlates with Theta power (Itthipuripat et al., 2013; Rutishauser et al., 2010). Our study, by establishing a negative coupling between Theta power and drift rate, reveals learning training's regulatory mechanism on cognitive resource demands: as multisensory categorical representations consolidate, individuals gradually reduce dependence on attention and working memory during classification decisions, thereby achieving more efficient feedforward information processing (enhanced drift rate). This finding provides new evidence for the "neural efficiency hypothesis" (Neubauer & Fink, 2009), indicating that multisensory learning may enhance decision system computational efficiency by optimizing neural oscillation patterns to reduce cognitive load.

Memory Retrieval Stage   The Late Positive Component (LPC), as a neural index of memory retrieval, shows amplitude enhancement closely associated with precise memory representation retrieval (Kwon et al., 2023; Sun et al., 2024). Research also links LPC to sustained attention and deep-level information processing (Gable & Harmon-Jones, 2013). Our study found that middle and late stages elicited LPC amplitude changes compared with early learning. Multiple regression further revealed that LPC significantly predicted correct-option bias shifts. At advanced cognitive stages, LPC dynamics reveal that individuals optimize decisions by accessing categorical representations in long-term memory systems for deep-level information processing. This process may reflect dynamic balancing mechanisms between neural resource reallocation and cognitive control. These results suggest that LPC may serve as an important neural indicator for memory retrieval and categorical decision-making in multisensory category learning.

Integrating ERP-DDM modeling results revealed that as learning progressed, the early perceptual system (P1) enhanced perceptual encoding specificity, thereby reducing downstream categorical feature discrimination load (N250-FSP). Simultaneously, activation patterns in the late long-term memory system

(LPC) indicated that individuals maintained optimal decision strategies by extracting effective features from long-term memory as learning advanced. The significant prediction of correct-choice bias by all three components provides a novel computational neural framework for understanding how learning experience promotes cognitive control optimization, revealing hierarchical interactive mechanisms among perceptual encoding, feature discrimination, and memory retrieval in decision dynamics.

Limitations and Future Directions

This study has several limitations that future research should address. First, six participants were excluded from final data analysis due to significant difficulties in category knowledge acquisition, indicating cognitive heterogeneity across individuals. Future studies should construct graded task difficulty systems to systematically examine how task difficulty parameters modulate multisensory category learning performance, thereby revealing dynamic associations between cognitive load and learning efficiency. Second, although the current paradigm effectively controlled irrelevant variables, its ecological validity could be enhanced. We plan to develop a VR-based multimodal interactive paradigm that simulates real learning scenarios while maintaining experimental control precision, simultaneously recording eye movements, gestures, and other multimodal behavioral data to deeply analyze cross-modal information integration mechanisms in dynamic contexts. Third, due to limited sampling range, the current study could not examine age effects. Future research should adopt cross-age longitudinal designs combined with fNIRS and other neuroimaging techniques to investigate developmental trajectories of multisensory category learning neural representations across the lifespan, constructing an integrated theoretical model incorporating cognitive development and neuroplasticity.

Conclusion

This study, using EEG and computational modeling methods, reveals the cognitive characteristics and neural mechanisms of multisensory category learning. Behavioral results demonstrate that as learning progressed, participants' information accumulation rates increased significantly, driving systematic improvements in task accuracy. Time-domain ERP analysis found that early perceptual component (P1) and late memory component (LPC) amplitudes increased with learning proficiency, while the N250-FSP complex associated with feature discrimination showed significant amplitude reduction.

Bayesian modeling based on the drift-diffusion model revealed a dual optimization mechanism: (1) enhanced information accumulation rates significantly correlated with reduced N250-FSP amplitudes and Theta band power; (2) starting point shifts toward prior experience were jointly predicted by P1, N250-FSP, and LPC.

These findings construct a three-stage theoretical framework of "perceptual

encoding-feature discrimination-memory retrieval": multisensory training optimizes the decision system hierarchically by reducing N250-FSP feature discrimination load while enhancing P1 perceptual representation and LPC memory retrieval efficiency. This study provides the first evidence for Theta oscillation's mediating role in linking neural activity with computational model parameters, offering a novel theoretical framework for understanding learning-induced cognitive resource reallocation.

References

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America, 102*(51), 18751-18756. doi:10.1073/pnas.0507704102

Bolam, J., Diaz, J. A., Andrews, M., Coats, R. O., Philiastides, M. G., Astill, S. L., & Delis, I. (2024). A drift diffusion model analysis of age-related impact on multisensory decision-making processes. *Scientific Report, 14*(1). doi:10.1038/s41598-024-65549-5

Cappe, C., Thelen, A., Romei, V., Thut, G., & Murray, M. M. (2012). Looming Signals Reveal Synergistic Principles of Multisensory Integration. *Journal of Neuroscience, 32*(4), 1171-1182. doi:10.1523/jneurosci.5517-11.2012

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Science, 18*(8), 414-421. doi:10.1016/j.tics.2014.04.012

Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueroa, C. M., Samanta, J., Sherman, S. J., & Frank, M. J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nature Neuroscience, 14*(11), 1462-1467. doi:10.1038/nn.2925

Clouter, A., Shapiro, K. L., & Hanslmayr, S. (2017). Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. *Current Biology, 27*(20), 3143-3148.e3146. doi:10.1016/j.cub.2017.09.001

Delorme, A., & Makeig, S. (2004). EEGLAB an open-source toolbox for analysis of single-trial EEG. *Journal of Neuroscience Methods, 134*, 9-21.

Faul, F., Erdfelder, E., & Lang, A., G. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences.* Behavior research methods, 39*(2), 175–191.

Folstein, J. R., Fuller, K., Howard, D., & DePatie, T. (2017). The effect of category learning on attentional modulation of visual cortex. *Neuropsychologia, 104*, 18-30. doi:10.1016/j.neuropsychologia.2017.07.025

Folstein, J. R., Monfared, S. S., & Maravel, T. (2017). The effect of category learning on visual attention and visual representation. *Psychophysiology,*

54(12), 1855-1871. doi:10.1111/psyp.12966

Folstein, J. R., Palmeri, T. J., & Gauthier, I. (2013). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cerebral Cortex, 23*(4), 814-823. doi:10.1093/cercor/bhs067

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, a. E. K. (2003). A Comparison of Primate Prefrontal and InferiorTemporal Cortices duringVisual Categorization. *The Journal of Neuroscience, 23*(12),

Gable, P. A., & Harmon-Jones, E. (2013). Does arousal per se account for the influence of appetitive stimuli on attentional scope positive potential? *Psychophysiology, 50*(4), https://doi.org/10.1111/psyp.12023

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Science, 10*(6), 278-285. doi:10.1016/j.tics.2006.04.008

Giard, M. H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience, 11*(5), 473-490. doi:10.1162/089892999563544

Goldstone, R. L. (1998). Perceptual Learning. *Annual Review of Psychology, 49*, 585-612.

Gui, P., Ku, Y., Li, L., Li, X., Bodner, M., Lenz, F. A., . . . Zhou, Y. D. (2017). Neural correlates of visuo-tactile crossmodal paired-associate learning and memory in humans. *Neuroscience, 362*, 181-195. doi:10.1016/j.neuroscience.2017.08.035

Herz, Damian M., Zavala, Baltazar A., Bogacz, R., & Brown, P. (2016). Neural Correlates of Decision Thresholds in the Human Subthalamic Nucleus. *Current Biology, 26*(7), 916-920. doi:10.1016/j.cub.2016.01.051

Hu, Z., Zhang, R., Zhang, Q., Liu, Q., & Li, H. (2012). Neural correlates of audiovisual integration of semantic category information. *Brain and Language, 121*(1), 70-75. doi:10.1016/j.bandl.2012.01.002

Itthipuripat, S., Wessel, J. R., & Aron, A. R. (2013). Frontal theta is a signature of successful working memory manipulation. *Experiment Brain Research, 224*(2), 255-262. doi:10.1007/s00221-012-3305-3

Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., Vanmeter, J., & Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron, 53*(6), 891-903. doi:10.1016/j.neuron.2007.02.015

Jiang, X., Chevillet, M. A., Rauschecker, J. P., & Riesenhuber, M. (2018). Training Humans to Categorize Monkey Calls: Auditory Feature- and Category-Selective Neural Tuning Changes. *Neuron, 98*(2), 405-416 e404. doi:10.1016/j.neuron.2018.03.014

Kawahara, H., & Matsui, H. (2003). "Auditory morphing based on an elastic perceptual distance metric in an interference-free time-frequency representation," in *International Conference on Acoustics, Speech, and Signal Processing, 2003* (Hong Kong: IEEE), 256–259

Keller, A. S., Payne, L., & Sekuler, R. (2017). Characterizing the roles of alpha and theta oscillations in multisensory attention. *Neuropsychologia, 99*, 48-63. doi:10.1016/j.neuropsychologia.2017.02.021

Kwon, S., Rugg, M. D., Wiegand, R., Curran, T., & Morcom, A. M. (2023). A meta-analysis of event-related potential correlates of recognition memory. *Psychonomic Bulletin & Review, 30*(6), 2083-2105. doi:10.3758/s13423-023-02309-y

Li, K.,Fu, Q., Fu, X. (2012). Cognitive and Neural Mechanisms of Probabilistic Category Learning. *Progress in Biochemistry and Biophysics, 39*(11), 1037-1044.

Li, Z., Lei, M., & Liu, Q. (2024). Cognitive mechanisms underlying the formation of offline representations in visual working memory. *Acta Psychologica Sinica, 56*(4), 412-420. doi:10.3724/sp.J.1041.2024.00412

Liu, Z., Zhang, Y., Ma, D., Xu, Q., & Seger, C. A. (2021). Differing effects of gain and loss feedback on rule-based and information-integration category learning. *Psychonomic Bulletin & Review, 28*(1), 274-282. doi:10.3758/s13423-020-01816-6

Mąka, S., Chrustowicz, M., & Okruszek, Ł. (2023). Can we dissociate hypervigilance to social threats from altered perceptual decision - making processes in lonely individuals? An exploration with Drift Diffusion Modeling and event-related potentials. *Psychophysiology, 60*(12). doi:10.1111/psyp.14406

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190. doi:10.1016/j.jneumeth.2007.03.024

Mück, M., Ohmann, K., Dummel, S., Mattes, A., Thesing, U., & Stahl, J. (2020). Face Perception and Narcissism: Variations of Event-Related Potential Components (P1 & N170) with Admiration and Rivalry. *Cognitive, Affective, & Behavioral Neuroscience, 20*(5), 1041-1055. doi:10.3758/s13415-020-00818-0

Neubauer, A., C.,, & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience and Biobehavioral Reviews, 33*, 1004–1023. doi:10.1016/j.neubiorev.2009.04.001

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 156869. doi:10.1155/2011/156869

Petrov, A. A. (2011). Category rating is based on prototypes and not instances: evidence from feedback-dependent context effects. *Journal of Experimental Psychology: Human Perception and Performance, 37*(2), 336-356.

doi:10.1037/a0021436

Roark, C. L., Lescht, E., Hampton Wray, A., & Chandrasekaran, B. (2023). Auditory and visual category learning in children and adults. *Developmental Psychology, 59*(5), 963-975. doi:10.1037/dev0001525

Roark, C. L., Paulon, G., Sarkar, A., & Chandrasekaran, B. (2021). Comparing perceptual category learning across modalities in the same individuals. *Psychonomic Bulletin & Review, 28*(3), 898-909. doi:10.1038/s13423-021-01878-0

Rutishauser, U., Ross, I. B., Mamelak, A. N., & Schuman, E. M. (2010). Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature, 464*(7290), 903-907. doi:10.1038/nature08860

Scholl, C. A., Jiang, X., Martin, J. G., & Riesenhuber, M. (2014). Time course of shape and category selectivity revealed by EEG rapid adaptation. *Journal of Cognitive Neuroscience, 26*(2), 408-421. doi:10.1162/jocn_a_{00477}

Scott, L., S„ Tanaka, J. W., Sheinberg, D., L„ & Curran, T. (2006). A Reevaluation of the Electrophysiological Correlates of Expert Object Processing. *Journal of Cognitive Neuroscience, 18*(9), 1453–1465.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2008). The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain Research, 1210*, 204-215. doi:10.1016/j.brainres.2008.02.054

Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. *Annual Review of Neuroscience, 33*, 203-219. doi:10.1146/annurev.neuro.051508.135546

Senkowski, D., Saint-Amour, D., Hofle, M., & Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage, 56*(4), 2200-2208. doi:10.1016/j.neuroimage.2011.03.075

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention The Colavita effect revisited.pdf. *Perception & Psychophysics, 69*(5), 673-686.

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., . . . Wallace, M. T. (2014). Identifying and quantifying multisensory integration: a tutorial review. *Brain Topography, 27*(6), 707-730. doi:10.1007/s10548-014-0365-7

Sun, J., Osth, A. F., & Feuerriegel, D. (2024). The late positive event-related potential component is time locked to the decision in recognition memory tasks. *Cortex, 176*, 194-208. doi:10.1016/j.cortex.2024.04.017

Tanaka, J., M., Curran, T., Porterfield, A., L. , & Collins, D. (2006). Activation of Preexisting and Acquired Face Representations: The N250 Event-related Potential as an Index of Face Familiarit. *Journal of Cognitive Neuroscience, 18*(9), 1488-1498. doi: 10.1162/jocn.2006.18.9.1488

Thelen, A., Cappe, C., & Murray, M. M. (2012). Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *Neuroimage, 62*(3), 1478-1488. doi:10.1016/j.neuroimage.2012.05.027

Thelen, A., Matusz, P. J., & Murray, M. M. (2014). Multisensory context portends object memory. *Current Biology, 24*(16), R734-735. doi:10.1016/j.cub.2014.06.040

Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage, 55*(3), 1208-1218. doi:10.1016/j.neuroimage.2010.12.068

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics, 7*. doi:10.3389/fninf.2013.00014

Wu, J., Li, Q., Fu, Q., Rose, M., & Jing, L. (2021). Multisensory Information Facilitates the Categorization of Untrained Stimuli. *Multisensory Research, 35*(1), 79-107. doi:10.1163/22134808-bja10061

Yang, X., Ying, C., Zhu, L., & Wang, WJ. (2024) The neural oscillations in delta-and theta-bands contribute to divided attention in audiovisual integration. *Perception, 53*(1), 44-60. doi: 10.1177/03010066231208539

Yuan, B., Wang, X., Yin, J., & Li, W. (2023). The role of cross-situational stimulus generalization in the formation of trust towards face: A perspective based on direct and observational learning. *Acta Psychologica Sinica, 55*(7). doi:10.3724/sp.J.1041.2023.01099

Zhou, G., Lane, G., Noto, T., Arabkheradmand, G., Gottfried, J. A., Schuele, S. U., . . . Zelano, C. (2019). Human olfactory-auditory integration requires phase synchrony between sensory cortices. *Nature Communication, 10*(1), 1168. doi:10.1038/s41467-019-09091-3

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.