Supplemental Materials

Model

The following introduces a prediction-based error-driven learning approach (adapts from the edl package Rij & Hoppe, 2021). The model utilizes a two-layer network, where the cue and outcome layers represent perceptual input and high-level expectations, respectively. Through learning, perceived cue i' (with apostrophe specifying cue presence) is discriminatively associated with expected or other outcome j, leading to an update $\Delta\omega_{i'j}$ in association weight $\omega_{i'j}$ at time step t. Note that throughout the Model section, t in the upper corner always denotes a time step rather than an exponent. This model incorporates an additional noise term ϵ, approximating a normal distribution (i.e., $log_e((1-x)/x)$), to account for trial-by-trial variations.

$$\omega_{i'j}^{(t+1)} = \omega_{i'j}^{(t)} + \Delta\omega_{i'j}^{(t)} + 0.001 \cdot \epsilon, \text{ for } i' \in I^{(t)} \tag{S1}$$

In this learning equation, the prediction error represents the difference between an outcome's maximum association weight λ_{max} (earlier conceptualized as the objective stimulus) and its current learned total association weight V_{total} across all learned cues at a given time step t. An additional scaling factor, α, is introduced to modulate the learning rate. In this case, however, the learning rate is not scaled ($\alpha = 1$).

$$\Delta\omega_{i'j}^{(t)} = \alpha \cdot \left(\lambda - V_{\text{total}}^{(t)}\right) \text{ where } V_{\text{total}}^{(t)} = \sum_{i \in I} \omega_{ij}^{(t)} \tag{S2}$$

While the currently learned association weight V_{total} remains below its maximum ($V_{total} < \lambda_{max}$ or 1), the associations of perceived cues at the current time step ($i' \in I^{(t)}$, where $i' \neq \varnothing$) are positively reinforced with the presented outcome j (where $j \neq \varnothing$) and negatively reinforced with absent outcomes ($j \in J$, where $j = \varnothing$). Discriminative learning is not applied when either all cues are absent (i.e., $i' = \varnothing$) or there is no prediction error for any outcome (i.e., blocking, where $V_{total} = \lambda_{max}$ or 1).

$$\Delta\omega_{i'j}^{(t)} = \begin{cases} \alpha \cdot \beta \cdot \left(1 - V_{\text{total}}^{(t)}\right) & \text{if } i' \neq \emptyset,\ j \neq \emptyset, \text{ and } V_{\text{total}} < 1 \\ \alpha \cdot \beta \cdot \left(0 - V_{\text{total}}^{(t)}\right) & \text{if } i' \neq \emptyset, \text{ and } j = \emptyset \\ 0 & \text{if } i' = \emptyset, \text{ or } V_{\text{total}} \geq 1 \end{cases} \tag{S3}$$

As mentioned, the error-driven learning model is atemporal and supervised, meaning it learns from labeled cue-outcome co-occurrences at each time step representing a stimulus presentation. Despite this framework, the model's learned cue-outcome associations can be applied either top-down or bottom-up to calculate cue or outcome activation (see Hoppe et al., 2022). Conventionally, the activation of a specific outcome at a specific time step ($A_{outcome}$) is dependent on learned associations with currently perceived cues from the bottom up. When anticipating a specific outcome, the model highlights the activation of informative perceptual cues over other shared or irrelevant ones from the top down (A_{cue}), leveraging the same learned associations.

This paper utilizes cue and outcome activation to simulate the cumulative neural responses observed in perceptual areas (e.g., temporal/occipital area) and expectation levels (e.g., frontal area). The approach is consistent with the sharpening view by demonstrating how top-down information enhances stimulus-specific perceptual responses. In terms of the facilitation view, however, prediction-based activation enhancement only incompletely addresses the information accumulation aspect. Activation at the perceptual cue and outcome expectation levels would further influence response latency, modulated by the activation itself (see Model 1), and additionally with off-task novelty seeking (see Model 2). In this paper, we refer to the result of information accumulation as activation (denoted by A) and the subsequent attenuation of this activation as neural response (denoted by $\mathcal{A}$). Note that the magnitude of neural responses and the value of response latency in the simulation study are arbitrary, which are nevertheless sufficient for modeling experience-related learning dynamics and demonstrating the direction of surprise-induced effects.

Model 1: Latency Attenuated

Neural response in the form of activation can be further transformed into its response latency, based on a retrieval latency equation (Anderson, 2007). A study by van der Velde et al. (2022) showed that this equation can directly correspond to the accumulation model proposed by Ratcliff & Smith (2004). The equation incorporates the natural e raised to a negative exponent of the continually updated activation A at time step t. This results in an exponential lowering of response latency γ as the activation increases. The response latency can be further scaled by (a) the initial response latency F, and/or (b) the rate of latency attenuation at the exponent f. In this simulation study, we are particularly interested in the exponent factor f (a direct scalar of A) that moderates the rate of latency attenuation.

$$\gamma^{(t)} = F \cdot e^{-f \cdot A^{(t)}} \tag{S4}$$

Due to the atemporal nature of the error-driven learning model, the latency-attenuated neural response $_{cue}$ to a specific stimulus at any given time step (epoch) is simply the product of its cue activation A_{cue} and its latency γ_{cue}. Additionally, the domain-specific feature of the cue further determines the spatial region associated with the response. For example, a perceived speech cue would elicit a response in the temporal area, while an image would elicit a response in the occipital area.

$$\mathcal{A}_{\text{cue}}^{(t)} = \gamma_{\text{cue}}^{(t)} \cdot A_{\text{cue}}^{(t)} \tag{S5}$$

The error-driven learning model can also derive outcome activation at the expectation level. Beyond attributing total cue-outcome associations to inform outcome activation (i.e., $A_{outcome}$, see Rij & Hoppe, 2021), this model considers the total latency-attenuated neural response of each cue (i.e., $_{cue}$) in anticipation of an outcome. In other words, outcome activation arises from integrating latency-attenuated perceptual responses $\sum_{i' \in I^{(t)}} {}_{cue}$ is modulated once again by response latency $\gamma_{outcome}$ at the expectation level. When latency attenuation is minimized (e.g., $f = 0$), neural response $\mathcal{A}$ converges to the activation A.

$$\begin{aligned} \mathcal{A}_{\text{outcome}}^{(t)} &= \gamma_{\text{outcome}}^{(t)} \cdot A_{\text{outcome}}^{(t)} \\ &= \gamma_{\text{outcome}}^{(t)} \cdot \sum_{i' \in I^{(t)}} \mathcal{A}_{\text{cue}}^{(t)} \end{aligned} \tag{S6}$$

Note that we only apply cue activation in EDL, using it alongside Equations 4 and 5 to calculate response latency and neural responses. The outcome activation from EDL is not applied in the current study; instead, it is computed via Equation 6 as the sum of perceptual neural responses. There are several reasons to adopt such an approach in terms of cognitive plausibility, particularly when considering the online processing of a task.

This process begins with stimulus-specific neural responses that have been sharpened by the system's history of learning from past prediction errors. As a specific task unfolds, these perceptual neural responses are actively enhanced within their respective sensory areas (e.g., the perception of a rattle sound). At this stage, activation within expectation-related frontal areas is driven by these real-time, ongoing perceptual signals rather than by previously learned associations alone. These dynamic sensory responses are then fed forward in real time to the frontal areas where the expected outcome is retrieved. The cumulative magnitude of these incoming signals determines the latency of the subsequent neural responses within expectation-related areas. Thus, even though the underlying EDL framework is atemporal, implementing Equation 6 allows the model to reflect this dynamic feedforward process.

Through this mechanism, our approach is compatible with predictive coding theory, where stimulus-specific neural activation at the perceptual areas is driven by top-down prediction errors from expectation-related areas (compatible with the sharpening view, as compared to the fatigue view). However, we do not differentiate whether such a prediction-based error detection mechanism needs to be conscious or not; we only consider that these expectation-related neural responses receive feedforward signals from the perceptual areas (as related to the facilitation view). Our approach slightly differs from theories that assume that expectation-related neural responses can only happen during conscious awareness.

Model 2: Latency+Novelty Attenuated

Infants are constantly exploring their environment during wakeful states. This means when latency attenuation frees up off-task temporal resources, it may lead to mind wandering or attention shifting toward novel activities. The latency equation (i.e.,) defines the maximum response time γ_{max} as 1 (note that in this context, the equation for response latency only applies to non-negative values, thus $\gamma_{cue/outcome} \in (e^{-1}, 1]$). As activation for either the cue or outcome gradually accumulates over time, the availability of additional time unrelated to the task (represented by $\gamma_{max} - \gamma_{cue/outcome}$) gradually increases. This extra off-task time allows for novelty-seeking neural responses to compete with

task-related neural responses. The model incorporates a constant level of activation for novelty seeking (represented by $A_{novelty} = 0.50$) that is modulated by the extra temporal resource. In this way, off-task novelty seeking now competes with and further attenuates on-task neural response.

$$\begin{cases} \mathcal{A}_{\text{cue}}^{(t)} = \gamma_{\text{cue}}^{(t)} \cdot A_{\text{cue}}^{(t)} - \left(\gamma_{max} - \gamma_{\text{cue}}^{(t)}\right) \cdot A_{\text{novelty}} \\ \mathcal{A}_{\text{outcome}}^{(t)} = \gamma_{\text{outcome}}^{(t)} \cdot A_{\text{outcome}}^{(t)} - \left(\gamma_{max} - \gamma_{\text{outcome}}^{(t)}\right) \cdot A_{\text{novelty}} \end{cases} \quad \text{(S7)}$$

Simulation Studies

The above-mentioned latency (Model 1) and latency-novelty (Model 2) attenuation models are applied to the learning of single-cue (Nakano et al., 2008) and multi-cue (Emberson et al., 2015; Kersey & Emberson, 2016) paradigms. The following simulation studies explore trial-by-trial changes in neural response by altering the latency attenuation factor ($f = 1$ or 3, Model 1) or further adding off-task novelty competition (Model 2).

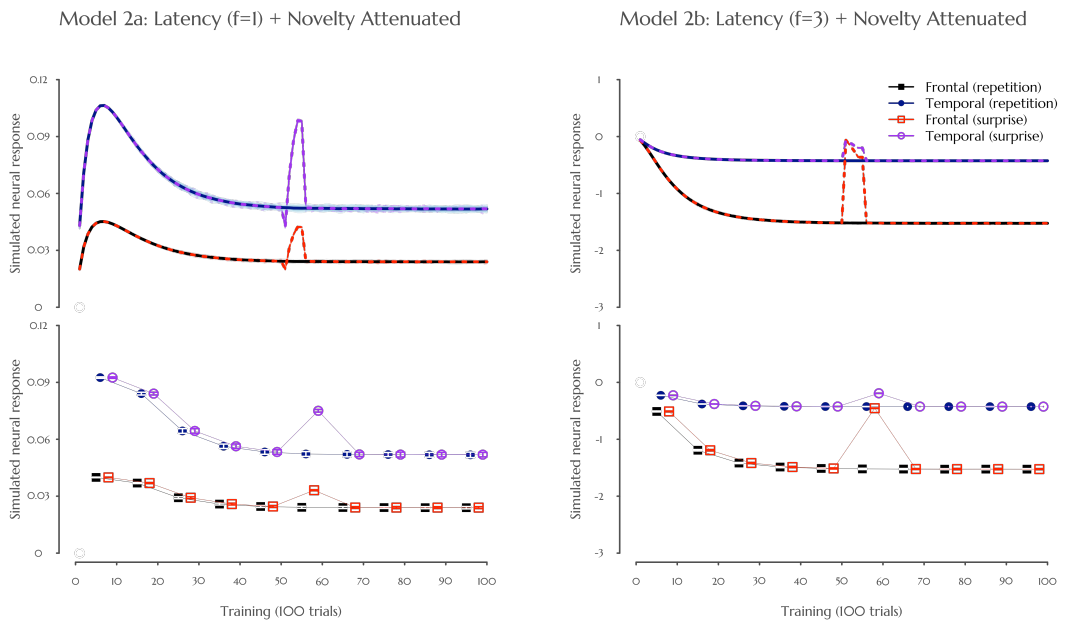


Figure S1. Simulated neural responses to the repetition of a single-cue stimulus (adapted from Nakano et al., 2008). Models 2a and 2b represent the same model with novelty attenuation, additional to respective default ($f = 1$) and higher ($f = 3$) latency attenuation. The top panel displays continuous changes in neural response over 100 trials. The bottom panel presents the same results binned into 10 blocks. Note that the 5th block of the surprise condition contains only the average responses from the repeated surprise stimulus. The x-axis represents time steps (top) or blocks (bottom), and the y-axis represents simulated neural response. In the top panel, solid lines indicate the repeated condition, whereas dashed lines represent the surprise condition. In the bottom panel, the 95% confidence intervals are indicated based on the error bars (negligible or null within the size of the points). Colors correspond to neural response regions associated with each condition: frontal lobe (black) and temporal lobe (blue) for the repetition condition, and frontal lobe (red) and temporal lobe (purple) for the surprise condition. Surprise trials occurred at time steps 51 and 55. Results are averaged across 50 independent model runs for both the repetition and surprise conditions.

Single-Cue Scenario

Model 2: Latency+Novelty Attenuated

The latency+novelty attenuation models 2a and 2b both exhibited repetition suppression and a corresponding surprise-induced enhancement (see Figure 1). While the simulation results align with Nakano et al. (2008), Model 2's inclusion of off-task novelty seeking primarily addresses wakeful exploration-exploitation (Hunter et al., 1983). Consequently, the simulation results are best interpreted in the context of repetition suppression during wakeful states (see repetition suppression among wakeful infants, Emberson et al., 2019). During wakeful states, the interpretation of the *f* parameter would slightly differ from that in the sleep state. Although infants may still integrate stimuli rapidly leading to more efficient responses, they can sustain attention, even passively, to the stimulus (e.g., Gomes et al., 2000). This implies that latency attenuation is smaller during wakeful states compared to sleep states (see Taga et al., 2018). Response attenuation is therefore likely attributable to off-task novelty-seeking rather than solely to latency attenuation. For example, electrophysiological studies have shown that repetition suppression among wakeful infants is associated with the reorientation of attention toward novel stimuli (Snyder & Keil, 2008). Among adults, moderate and strong neural attenuation has been observed during passive task performance (Oades & Dittmann-Balcar, 1995) and forced attention shifts (Woldorff et al., 1998), respectively.

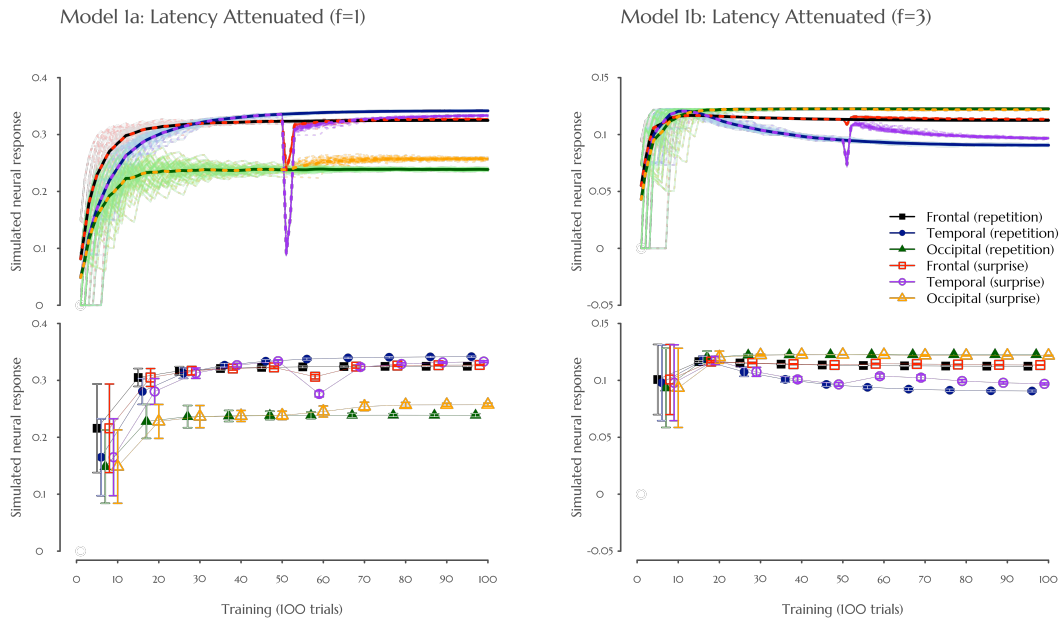


Figure S2. Simulated neural responses to the repetition of a multi-cue stimulus (adapted from Emberson et al., 2015; Kersey & Emberson, 2016). Models 1a and 1b represent the same model with default ($f = 1$) and higher ($f = 3$) latency attenuation, respectively. The top panel displays continuous changes in neural response over 100 trials. The bottom panel presents the same results binned into 10 blocks. Note that the 5th block of the surprise condition contains only the average responses from the repeated surprise stimulus. The x-axis represents time steps (top) or blocks (bottom), and the y-axis represents simulated neural response. In the top panel, solid lines indicate the repeated condition, whereas dashed lines represent the surprise condition. In the bottom panel, the 95% confidence intervals are indicated based on the error bars. Colors correspond to neural response regions associated with each condition: frontal lobe (black), temporal lobe (blue), and occipital lobe (green) for the repetition condition, and frontal lobe (red), temporal lobe (purple), and occipital lobe (orange) for the surprise condition. Surprise trials occurred at time steps 51 and 52. Results are averaged across 50 independent model runs for both the repetition and surprise conditions.

Prolonged measuring of experience-related neural dynamics in infants is challenging due to infants easily waning attention during repetitive tasks. Consequently, the following primarily considers mismatch-related electrophysiological literature, assuming the initial stimulus presentation is similar to a surprise-induced response (in EDL, the first surprising stimulus and the first stimulus occurrence are equivalent for simple stimuli with unique cues, but this equivalence does not hold for complex stimuli that share cues). The results of Model 2a are consistent with corresponding neural responses (i.e., N1- and additional MMN-like responses) observed in wakeful infants (Cheng et al., 2013; Fellman et al., 2004; He et al., 2007; Kushnerenko et al., 2002). Other studies have shown that school-aged children demonstrate passive mismatch responses even while engaged in alternative tasks like reading or watching videos, with the amplitude of these enhanced mismatch responses comparable to those elicited by active attention (see Gomes et al., 2000). In our model, the enhanced surprise-induced response is interpreted as rapid learning of the equally simple, novel single-cue stimulus. The overall enhancement of the novel stimulus resets novelty attenuation, similar to the response to the first instances of the repeated stimulus. Furthermore, the response to a novel stimulus is dynamic and influenced by learning. Notably, the initial response to a novel stimulus is slightly reduced compared to the incomplete suppression caused by off-task competition, before a pronounced enhancement emerges (see upper panel of Model 2a results,).

Lastly, the results of Model 2b are based on an implausible parameter setting that incorporates rapid latency attenuation (i.e., $f = 3$). In other words, this model exhibits high latency attenuation during wakefulness, as if neural responses diminished rapidly as in the "sleep" model (the same f as in Model 1b). A pathological increase in auditory N1 (and P2) latency attenuation has been detected in dyslexic adults compared to controls using simple tones (Jaffe-Dax et al., 2017). Moreover, the attenuation rate in Jaffe-Dax et al. (2017) is also moderated by the inter-stimulus interval, consistent with other findings (Loizides et al., 2015; Wang et al., 2008). Last but not least, this atypical attenuation rate in dyslexics has been confirmed across various brain sites (e.g., both frontal and central regions Jaffe-Dax et al., 2017), potentially suggesting a prediction-based interaction between perception and expectation.

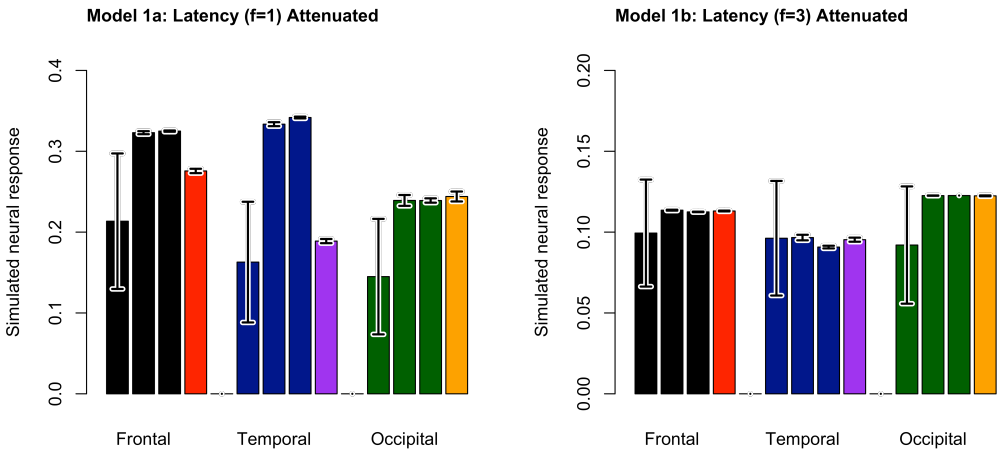


Figure S3. To better illustrate the overall experience-related neural dynamics, bar plots are provided for each model, showing neural responses at the 1st, 5th, and 10th blocks (each averaged across 10 repetitions), followed by the surprise-induced block (averaged across 2 repetitions) from left to right. The error bars indicate the 95% confidence intervals. Color coding for repeated and surprise stimulus for each cortical area is consistent with Figure 2.

		Model 1a		*Model 1b*	
		beta	*t*	*beta*	*t*
temporal	*x*	*0.48*	*65.8*	*-0.28*	*30.1*
occipital	*x*	*0.34*	*38.1*	*0.25*	*27.1*
frontal	*x*	*0.36*	*42.0*	*9.17*	*9.3*
temporal	*x*	*0.48*	*97.1*	*-0.28*	*30.2*
	x^2	*-0.38*	*76.4*	*-4.00*	*4.4*
occipital	*x*	*0.34*	*45.5*	*0.25*	*29.9*
	x^2	*-0.34*	*46.2*	*-0.28*	*33.0*
frontal	*x*	*0.36*	*51.0*	*9.17*	*9.7*
	x^2	*-0.34*	*48.4*	*-0.95*	*19.0*

Table S1. Mixed-effects regressions were conducted on trial-by-trial responses simulated by Models 1a and 1b, separately for each cortical location. Statistical test results presented above the midline are based solely on a linear (trial) regressor, while those below incorporate both linear (trial) and quadratic (trial squared) regressors. All *p* values < 0.001.

Multi-Cue Scenario

Model 1: Latency Attenuated

For latency-attenuated Model 1a results, linear regression indicated consistently positive linear trends (trial) in trial-by-trial neural responses across all cortical areas. Additionally, quadratic regression revealed further negative quadratic trends (trial squared) after accounting for the corresponding linear trend (see Table 1, Figure 2). Quadratic regression models outperformed linear models in all cortical areas (deviance of linear regression subtracted from quadratic regression, i.e, χ^2: temporal = 70.76, occipital = 1046.10, frontal = 447.95; all $p < 0.001$). These findings align with Kersey & Emberson (2016). However, subsequent block-wise analyses demonstrated only repetition enhancement without suppression in all cortical areas (see Figure 3). This suggests that the quadratic trends observed in Model 1a primarily reflect repetition enhancement, which exhibits quadratic rather than linear patterns. Block-wise analyses also demonstrated surprise-induced suppressed temporal responses, with no changes in occipital response, consistent with Emberson et al., (2015). These results are counterintuitive given that the sound was presented in the surprise stimulus while the smiley image was absent. However, they can be interpreted within a prediction-based framework in terms of the informativeness of the perceptual signal.

On the other hand, the more rapid attenuation rate of latency-attenuated Model 1b results indicated a negative linear trend (by trial) in simulated temporal responses. A large positive linear trend was also revealed for simulated frontal responses, deviating from other coefficients. Furthermore, block-wise analyses revealed no evidence of surprise-induced repetition suppression in temporal responses. Taken together, the simulated results of Model 1b diverge from empirical findings

(Emberson et al., 2015; Kersey & Emberson, 2016), suggesting that a rapid latency attenuation resembling sleep-state dynamics is implausible in modeling experience-related neural dynamics in typically developing infants.

Sensitivity Analyses (on Novelty Attenuation)

The novelty-related activation value is varied from 0.30 to 0.70, in addition to the value ($A_{novelty} = 0.50$) used in the manuscript. These simulations specifically apply to the wakeful models (Models 2a) with a default rate of latency change ($f = 1$). Note that it is already mentioned in the manuscript that a higher rate of latency change during wakefulness is related to pathological neural responses. We first examine the simulations regarding Nakano et al. (2008). Model results exhibit the repetition-related and surprise-induced neural responses reported in wakeful infants (Emberson et al., 2019) only when the novelty-related activation value exceeded 0.50 (also see Figure 1 in the Supplemental Materials). When the novelty-related activation value is lower (less than 0.50), the simulation findings resemble the results of Model 1a (see Figure 5 in the manuscript). This suggests that the model's main findings require a relatively higher novelty-related activation value.

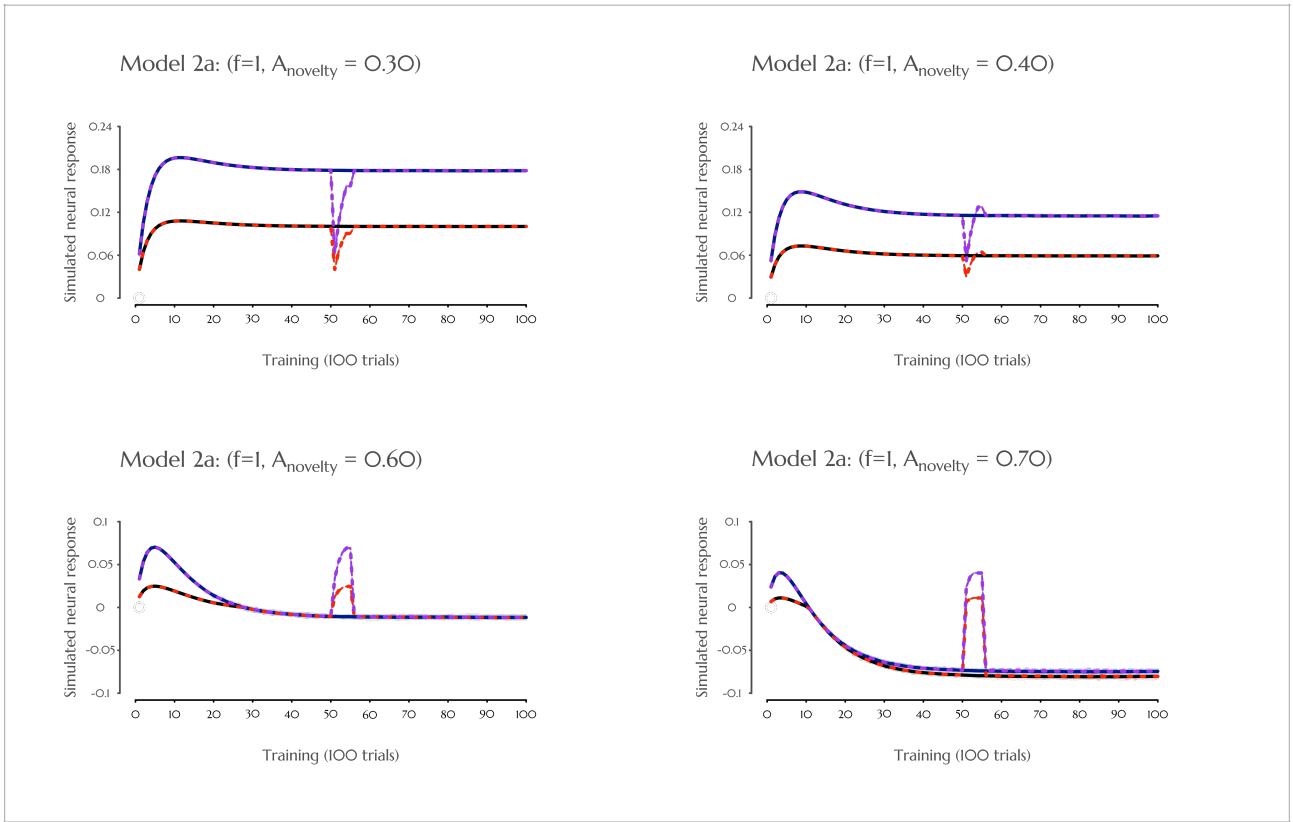


Figure S4. Simulations re. Nakano et al., (2008) where novelty attenuation is moderated. Models 2a represent the model with novelty attenuation with default latency attenuation ($f = 1$). Each figure displays continuous changes in neural response over 100 trials. The x-axis represents time steps, and the y-axis represents simulated neural response. Solid lines indicate the repeated condition, whereas dashed lines represent the surprise condition. Colors correspond to neural response regions associated with each condition: frontal lobe (black) and temporal lobe (blue) for the repetition condition, and frontal lobe (red) and temporal lobe (purple) for the surprise condition. Surprise trials occurred at time steps 51 and 55. Results are averaged across 50 independent model runs for both the repetition and surprise conditions.

We then examine the simulations regarding Kersey and Emberson (2016). Models also exhibit repetition-related and surprise-induced neural responses reported in wakeful infants only when the novelty-related activation value exceeded 0.50 (also see Figure 6 in the manuscript). When the novelty-related activation value is lower (less than 0.50), the simulation findings resemble the results of Model 1a (see Figure 2 in the Supplemental Materials). This likewise suggests that the model's main findings require a relatively higher novelty-related activation value.

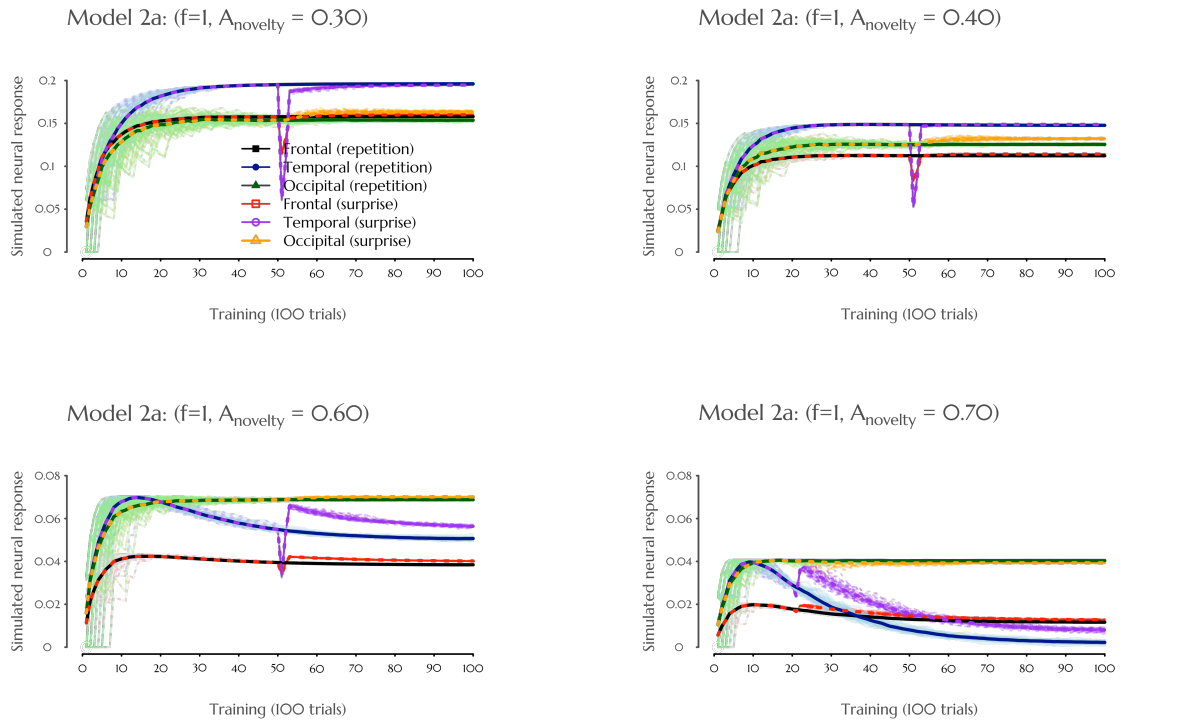


Figure S5. Simulations re. Kersey and Emberson, (2016) where novelty attenuation is moderated. Models 2a represent the model with novelty attenuation with default latency attenuation ($f = 1$). Each figure displays continuous changes in neural response over 100 trials. The x-axis represents time steps, and the y-axis represents simulated neural response. Solid lines indicate the repeated condition, whereas dashed lines represent the surprise condition. Colors correspond to neural response regions associated with each condition: frontal lobe (black), temporal lobe (blue), and occipital lobe (green) for the repetition condition, and frontal lobe (red), temporal lobe (purple), and occipital lobe (orange) for the surprise condition. Surprise trials occurred at time steps 51 and 52. Results are averaged across 50 independent model runs for both the repetition and surprise conditions.

Taken together, our sensitivity analyses suggest that the model's main findings remain robust when higher novelty-related activation values are applied (specifically, greater than 0.50 for both tasks). Theoretically, a higher novelty-related activation value is plausible for wakeful infants, who constantly explore their environment, quickly exhausting one activity before transitioning to an alternative.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* (pp. 237–248). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195324259.003.0006

Cheng, Y.-Y., Wu, H.-C., Tzeng, Y.-L., Yang, M.-T., Zhao, L.-L., & Lee, C.-Y. (2013). The development of mismatch responses to mandarin lexical tones in early infancy. *Developmental Neuropsychology*, 38(5), 281–300. https://doi.org/10.1080/87565641.2013.799672

Emberson, L. L., Boldin, A. M., Robertson, C. E., Cannon, G., & Aslin, R. N. (2019). Expectation affects neural repetition suppression in infancy. *Developmental Cognitive Neuroscience*, 37, 100597. https://doi.org/10.1016/j.dcn.2018.11.001 *(Note: Cited in-text as Emberson et al., 2019)*

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proceedings of the National Academy of Sciences*, 112(31), 9585–9590. https://doi.org/10.1073/pnas.1510343112

Fellman, V., Kushnerenko, E., Mikkola, K., Ceponiene, R., Leipälä, J., & Näätänen, R. (2004). Atypical auditory event-related potentials in preterm infants during the first year of life: A possible sign of cognitive dysfunction? *Pediatric Research*, 56(2), 291–297. https://doi.org/10.1203/01.pdr.0000132750.97066.b9

Gomes, H., Molholm, S., Ritter, W., Kurtzberg, D., Cowan, N., & Vaughan, H. G. (2000). Mismatch negativity in children and adults, and effects of an attended task. *Psychophysiology*, 37(6), 807–816. https://doi.org/10.1017/s0048577200990632

He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, 19(5), 878–892. https://doi.org/10.1162/jocn.2007.19.5.878

Hoppe, D. B., Hendriks, P., Ramscar, M., & Rij, J. van. (2022). An exploration of error-driven learning in simple two-layer networks from a discriminative learning perspective. *Behavior Research Methods*, 54(5), 2221–2251. https://doi.org/10.3758/s13428-021-01711-5

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, 19(3), 338–352. https://doi.org/10.1037//0012-1649.19.3.338

Jaffe-Dax, S., Frenkel, O., & Ahissar, M. (2017). Dyslexics' faster decay of implicit memory for sounds and words is manifested in their shorter neural adaptation. *eLife*, 6. https://doi.org/10.7554/elife.20557

Kersey, A. J., & Emberson, L. L. (2016). Tracing trajectories of audio-visual learning in the infant brain. *Developmental Science*, 20(6). https://doi.org/10.1111/desc.12480

Kushnerenko, E., Čeponienė, R., Balan, P., Fellman, V., & Näätänen, R. (2002). Maturation of the auditory change detection response in infants: A longitudinal ERP study. *NeuroReport*, 13(15), 1843–1848. https://doi.org/10.1097/00001756-200210280-00002 *(Note: Encoding errors resolved)*

Loizides, C., Achilleos, A., Iannetti, G. D., & Mitsis, G. D. (2015). Assessment of nonlinear interactions in event-related potentials elicited by stimuli presented at short interstimulus intervals

using single-trial data. *Journal of Neurophysiology*, 113(10), 3623–3633. https://doi.org/10.1152/jn.00523.2014

Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2008). Prefrontal cortical involvement in young infants' analysis of novelty. *Cerebral Cortex*, 19(2), 455–463. https://doi.org/10.1093/cercor/bhn096

Oades, R. D., & Dittmann-Balcar, A. (1995). Mismatch negativity (MMN) is altered by directing attention. *NeuroReport*, 6(8), 1187–1190. https://doi.org/10.1097/00001756-199505300-00028

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111(2), 333–367. https://doi.org/10.1037/0033-295x.111.2.333

Rij, J. van, & Hoppe, D. (2021). *Edl: Toolbox for error-driven learning simulations with two-layer networks*. In CRAN: Contributed Packages. The R Foundation. https://doi.org/10.32614/cran.package.edl

Snyder, K. A., & Keil, A. (2008). Repetition suppression of induced gamma activity predicts enhanced orienting toward a novel stimulus in 6-month-old infants. *Journal of Cognitive Neurocience*, 20(12), 2137–2152. https://doi.org/10.1162/jocn.2008.20149

Taga, G., Watanabe, H., & Homae, F. (2018). Developmental changes in cortical sensory processing during wakefulness and sleep. *NeuroImage*, 178, 519–530. https://doi.org/10.1016/j.neuroimage.2018.05.075

Velde, M. van der, Sense, F., Borst, J. P., Maanen, L. van, & Rijn, H. van. (2022). Capturing dynamic performance in a cognitive model: Estimating ACT-r memory parameters with the linear ballistic accumulator. *Topics in Cognitive Science*, 14(4), 889–903. https://doi.org/10.1111/tops.12614 *(Note: Cited in-text as van der Velde et al., 2022)*

Wang, A. L., Mouraux, A., Liang, M., & Iannetti, G. D. (2008). The enhancement of the n1 wave elicited by sensory stimuli presented at very short inter-stimulus intervals is a general feature across sensory systems. *PLoS ONE*, 3(12), e3929. https://doi.org/10.1371/journal.pone.0003929

Woldorff, M. G., Hillyard, S. A., Gallen, C. C., Hampson, S. R., & Bloom, F. E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology*, 35(3), 283–292. https://doi.org/10.1017/s0048577298961601