RESEARCH ARTICLE

WILEY

EEG mismatch responses in a multimodal roving stimulus paradigm provide evidence for probabilistic inference across audition, somatosensation, and vision

Miro Grundei[1,2] 🔵 | Pia Schröder[1] | Sam Gijsen[1,2] | Felix Blankenburg[1,2]

[1]Neurocomputation and Neuroimaging Unit, Freie Universität Berlin, Berlin, 14195, Germany

[2]Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, 10117, Germany

Correspondence
Miro Grundei, Neurocomputation and Neuroimaging Unit, Freie Universität Berlin, 14195 Berlin, Germany.
Email: m.grundei@fu-berlin.de

Abstract

The human brain is constantly subjected to a multimodal stream of probabilistic sensory inputs. Electroencephalography (EEG) signatures, such as the mismatch negativity (MMN) and the P3, can give valuable insight into neuronal probabilistic inference. Although reported for different modalities, mismatch responses have largely been studied in isolation, with a strong focus on the auditory MMN. To investigate the extent to which early and late mismatch responses across modalities represent comparable signatures of uni- and cross-modal probabilistic inference in the hierarchically structured cortex, we recorded EEG from 32 participants undergoing a novel tri-modal roving stimulus paradigm. The employed sequences consisted of high and low intensity stimuli in the auditory, somatosensory and visual modalities and were governed by unimodal transition probabilities and cross-modal conditional dependencies. We found modality specific signatures of MMN ($\sim$100–200 ms) in all three modalities, which were source localized to the respective sensory cortices and shared right lateralized prefrontal sources. Additionally, we identified a cross-modal signature of mismatch processing in the P3a time range ($\sim$300–350 ms), for which a common network with frontal dominance was found. Across modalities, the mismatch responses showed highly comparable parametric effects of stimulus train length, which were driven by standard and deviant response modulations in opposite directions. Strikingly, P3a responses across modalities were increased for mispredicted stimuli with low cross-modal conditional probability, suggesting sensitivity to multimodal (global) predictive sequence properties. Finally, model comparisons indicated that the observed single trial dynamics were best captured by Bayesian learning models tracking unimodal stimulus transitions as well as cross-modal conditional dependencies.

KEYWORDS
Bayesian inference, cross-modal, mismatch negativity, multisensory, P3, predictive processing, surprise

1 | INTRODUCTION

Humans inhabit a highly structured environment governed by complex regularities. The brain is subjected to such environmental regularities by a multimodal stream of sensory inputs ultimately constructing a perceptual representation of the world. The sensory system is thought to capitalize on statistical regularities to efficiently guide interaction with the world enabling anticipation and rapid detection of sensory changes (Bregman, 1994; Dehaene et al., 2015; Friston, 2005; Frost et al., 2015; Gregory, 1980; Winkler et al., 2009).

Neuronal responses to deviations from sensory regularities can be valuable windows into the brain's processing of statistical properties of the environment and corresponding sensory predictions. The presentation of rare deviant sounds within a sequence of repeating standard sounds induces well known mismatch responses (MMRs) that can be recorded with electroencephalography (EEG), such as the mismatch negativity (MMN; Naatanen et al., 1978; Naatanen et al., 2007) and the P3 (or P300; Polich, 2007; Squires et al., 1975; Sutton et al., 1965). The MMN is defined as a negative EEG component resulting from subtraction of standard from deviant trials between ~100 and 200 ms poststimulus. Although the MMN has primarily been researched in the auditory modality, similar early mismatch components have been reported for other sensory modalities, including the visual (Kimura et al., 2011; Pazo-Alvarez et al., 2003; Stefanics et al., 2014) and, to a lesser extent, the somatosensory modality (Andersen & Lundqvist, 2019; Hu et al., 2013; Kekoni et al., 1997). The P3 is a later positive going component in response to novelty between 200 and 600 ms around central electrodes, which has been described for the auditory, somatosensory, and visual modalities and is known for its modality independent characteristics (Escera et al., 2000; Friedman et al., 2001; Knight & Scabini, 1998; Polich, 2007; Schroger, 1996).

Despite being one of the most well-studied EEG components, the neuronal generation of the MMN remains subject of ongoing debate (Garrido, Kilner, Stephan, & Friston, 2009; May & Tiitinen, 2010; Naatanen et al., 2005). Two prominent but opposing accounts cast the MMN as adaptation-based or memory-based, respectively. Adaptation-based accounts argue that the observed differences between standard and deviant responses primarily result from neuronal attenuation leading to stimulus specific adaptation (SSA; Jaaskelainen et al., 2004; May et al., 1999). In animals, SSA has been shown to result in response patterns similar to the MMN (Ulanovsky et al., 2003; Ulanovsky et al., 2004) and simulation work suggests that different types of MMN-like responses can be reproduced by pure adaptation models (May & Tiitinen, 2010). However, it remains unclear if the full range of MMN characteristics can be explained by adaptation alone (Fitzgerald & Todd, 2020; Garrido, Kilner, Stephan, & Friston, 2009; Wacongne et al., 2012). The memory-based view, on the other hand, suggests that the MMN is a marker of change detection based on sensory memory trace formation (Näätänen, 1990; Naatanen et al., 2005; Naatanen & Näätänen, 1992). The memory trace stores local information on stimulus regularity and compares it to incoming sensory inputs that may signal changes in the current sensory stream.

While largely neglected by previous interpretations of the MMN, it is becoming increasingly clear that key empirical features of MMRs concern stimulus predictability rather than stimulus change per se. The MMN has been reported in response to abstract rule violations (Paavilainen, 2013), unexpected stimulus repetitions (Alain et al., 1994; Horvath & Winkler, 2004; Macdonald & Campbell, 2011) and unexpected stimulus omissions (Heilbron & Chait, 2018; Hughes et al., 2001; Salisbury, 2012; Wacongne et al., 2011; Yabe et al., 1997). Similar characteristics have been reported for P3 MMRs (Duncan et al., 2009; Prete et al., 2022) and both MMN and P3 responses have been shown to increase for unexpected compared to expected deviants (Schroger et al., 2015; Sussman, 2005; Sussman et al., 1998). Insights concerning the predictive nature of MMRs have led to further development of the memory-based account of MMN generation into the model-adjustment hypothesis (Winkler, 2007). This view assumes a perceptual model that is informed by previous stimulus exposure and continually predicts incoming sensory inputs. The model is updated whenever inputs diverge from current predictions, and the MMN is hypothesized to constitute a marker of such divergence.

The model-adjustment hypothesis is in line with the increasingly influential view that the brain is engaging in perceptual inference to anticipate future sensory inputs (Friston, 2005; Gregory, 1980; Von Helmholtz, 1867). Related theories regard the brain as an inference engine and come with neuronal implementation schemes that accomplish probabilistic (Bayesian) inference in a neurologically plausible manner (Bastos et al., 2012; Friston, 2005, 2010). Process theories such as predictive coding assume that the brain maintains a generative model of its environment which is continuously updated by comparing incoming sensory information with model predictions on different levels of hierarchical cortical organization (Friston, 2005, 2010; Rao & Ballard, 1999; Winkler & Czigler, 2012). Differential influences of SSA and change detection on the MMN are proposed to result from the same underlying process of prediction error minimization, mediated by different post-synaptic changes to (predicted) sensory inputs (Auksztulewicz & Friston, 2016; Garrido et al., 2008; Garrido, Kilner, Kiebel, & Friston, 2009). As such, the theory has the potential to unify previously opposing theories of MMN generation (Garrido et al., 2008; Garrido, Kilner, Kiebel, & Friston, 2009; Garrido, Kilner, Stephan, & Friston, 2009) while accounting for its key empirical features (Heilbron & Chait, 2018; Wacongne et al., 2012).

With regard to the proposed universal nature of predictive accounts of brain function, reports of comparable MMRs across different modalities are of particular interest. So far, mismatch signals have been primarily studied in isolation, with a strong focus on the auditory system. However, key properties of the auditory MMN, such as omission responses and modulations by predictability, have also been reported for the visual (Czigler et al., 2006; Kok et al., 2014) and the somatosensory MMN (Andersen & Lundqvist, 2019; Naeije et al., 2018), and modeling studies in all three modalities suggest that

MMRs may reflect signatures of Bayesian learning (BL; Gijsen et al., 2021; Lieder et al., 2013; Maheu et al., 2019; Ostwald et al., 2012; Stefanics et al., 2018). While studies directly investigating mismatch signals in response to multimodal sensory inputs are rare, previous research indicates a ubiquitous role for cross-modal probabilistic learning. The brain tends to automatically integrate auditory, somatosensory, and visual stimuli during sequence processing (Bresciani et al., 2006, 2008; Frost et al., 2015) and cross-modal perceptual associations can influence statistical learning of sequence regularities (Andric et al., 2017; Parmentier et al., 2011), modulate MMRs (Besle et al., 2005; Butler et al., 2012; Friedel et al., 2020; Kiat, 2018; Zhao et al., 2015) and influence subsequent unimodal processing in various ways (Shams et al., 2011). Recent advances in modeling Bayesian causal inference suggest that the main computational stages of multimodal inference evolve along a multisensory hierarchy involving early sensory segregation followed by mid-latency sensory fusion and late Bayesian causal inference (Cao et al., 2019; Rohe et al., 2019; Rohe & Noppeney, 2015). However, the extent to which the MMN and P3 reflect these stages and should be considered sensory specific signatures of regularity violation or the result of modality independent computations in an underlying predictive network is not fully understood.

The current study aimed to investigate the commonalities and differences between MMRs in different modalities in a single experiment and to elucidate in how far they reflect local, unimodal or global, cross-modal computations. To this end, we employed a roving stimulus paradigm, in which auditory, somatosensory, and visual stimuli were simultaneously presented in a probabilistic tri-modal stimulus stream.

Typically, MMRs are studied with the oddball paradigm, in which rarely presented "oddball" stimuli deviate from frequently presented standard stimuli in some physical feature, such as sound pitch or stimulus intensity. The roving stimulus paradigm, on the other hand, defines deviants and standards in terms of their local sequence position, while the frequency of occurrence of their stimulus features across the sequence is equal (Baldeweg et al., 2004; Cowan et al., 1993). The deviant is defined as the first stimulus that breaks a train of repeating (standard) stimuli. With repetition, the deviant subsequently becomes the new standard, defining a train of stimulus repetitions. Thus, the roving stimulus paradigm is an excellent tool to experimentally induce MMRs, while controlling for differences in physical stimulus features.

Based on a probabilistic model, we generated sequences of high and low intensity stimuli that were governed by unimodal transition probabilities as well as cross-modal conditional dependencies. This allowed us to test to what extent early and late MMRs are sensitive to local and global violations of statistical regularities and to draw conclusions regarding their potential role in cross-modal hierarchical inference. Specifically, we extracted the MMN and P3 MMRs for each modality and investigated their modality specific and modality general response properties regarding stimulus repetition and change, as well as their sensitivity to cross-modal predictive information. Further, we used source localization to investigate modality specific and modality

general neuronal generators of MMRs. Finally, we complemented our average-based analyses with single-trial modeling to investigate if signatures of unimodal and cross-modal Bayesian inference can account for trial-to-trial fluctuations in the MMN and P3 amplitudes.

2 | MATERIALS AND METHODS

Participants underwent a novel multimodal version of the roving stimulus paradigm. Our paradigm, depicted in Figure 1, consisted of simultaneously presented auditory (A), somatosensory (S), and visual (V) stimuli, which each alternated between two different intensity levels ("low" and "high"). The tri-modal stimulus sequences originated from a single probabilistic model (described in Section 2.3), resulting in different combinations of low and high stimuli across the three modalities in each trial.

2.1 | Participants

Thirty-four healthy volunteers (19–43 years old, mean age: 26, 22 females, all right-handed), recruited from the student body of the Freie Universität Berlin and the general public, participated for monetary compensation or an equivalent in course credit. The study was approved by the local ethics committee of the Freie Universität Berlin and written informed consent was obtained from all participants prior to the experiment.

2.2 | Experimental setup

Each trial consisted of three bilateral stimuli (A, S, and V) that were presented simultaneously by triggering three instantaneous outputs of a data acquisition card (National Instruments Corporation, Austin, Texas, USA) every 1150 ms (inter-stimulus interval).

Auditory stimuli were presented via in-ear headphones (JBL, Los Angeles, California, USA) to both ears and consisted of sinusoidal waves of 500 Hz and 100 ms duration that were modulated by two different amplitudes. The amplitudes were individually adjusted with the participants to obtain two clearly distinguishable intensities (mean of the low intensity stimulus: 81.43 ± 1.22 dB; mean of the high intensity stimulus: 93.02 ± 0.98 dB).

Somatosensory stimuli were administered with two DS5 isolated bipolar constant current stimulators (Digitimer Limited, Welwyn Garden City, Hertfordshire, UK) via adhesive electrodes (GVB-geliMED GmbH, Bad Segeberg, Germany) attached to the wrists of both arms. The stimuli consisted of electrical rectangular pulses of 0.2 ms duration. To account for interpersonal differences in sensory thresholds, the two intensity levels used in the experiment were determined on an individual basis. The low intensity level (mean: 3.97 ± 0.84 mA) was set in proximity to the detection threshold yet high enough to be clearly perceivable (and judged to be the same intensity on both sides). The high intensity level (mean: 6.47 ± 1.33 mA) was determined
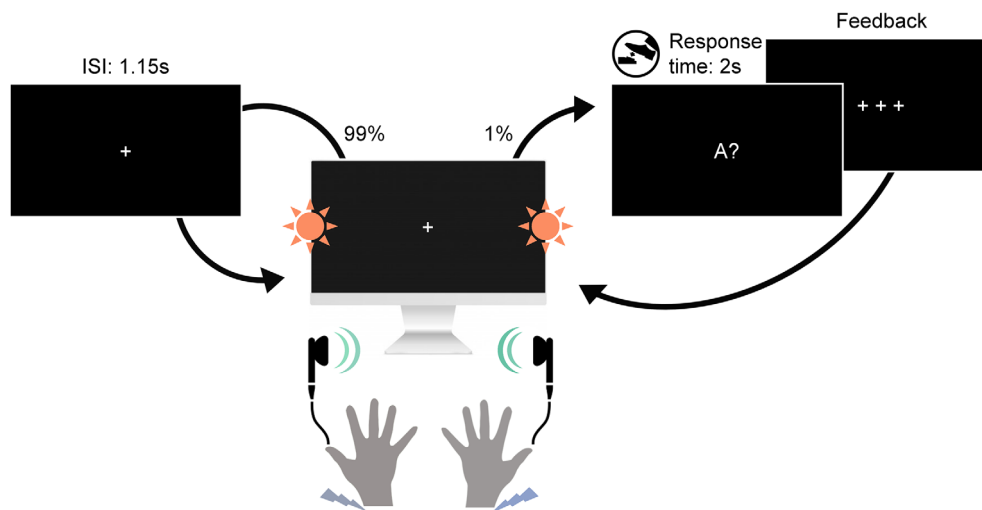
FIGURE 1   Experimental paradigm. Participants were seated in front of a screen and received sequences of simultaneously presented bilateral auditory beep stimuli (green), somatosensory electrical pulse stimuli (purple) and visual flash stimuli (orange) each at either low or high intensity. On consecutive trials, stimuli within each modality either repeated the previous stimulus intensity of that modality (standard) or alternated to the other intensity (deviant). This created tri-modal roving stimulus sequences, where the repetition/alternation probability in each modality was determined by a single probabilistic model (see Section 2.3). In 1% of trials (catch trials) the fixation cross changed to one of the three letters A, T, or V, interrupting the stimulus sequence. The letter prompted participants to indicate whether the last auditory (letter A), somatosensory (letter T for "tactile") or visual (letter V) stimulus, respectively, was of high or low intensity. Responses were given with a left or right foot pedal press using the right foot.

for each participant to be easily distinguishable from the low intensity level yet remaining non-painful and below the motor threshold.

Visual stimuli were presented via light emitting diodes (LEDs) and transmitted through optical fiber cables mounted vertically centered to both sides of a monitor. The visual flashes consisted of rectangular waves of 100 ms duration that were modulated by two different amplitudes (low intensity stimulus: 2.65 V; high intensity stimulus: 10 V) that were determined to be clearly perceivable and distinguishable prior to the experiment. Participants were seated at a distance of about 60 cm to the screen such that the LED's were placed within the visual field at a visual angle of about 67°.

In each of six experimental runs of 11.5 min, a sequence of 600 stimulus combinations was presented. To ensure that participants maintained attention throughout the experiment and to encourage monitoring of all three stimulation modalities, participants were instructed to respond to occasional catch trials (target questions) via foot pedals. In six trials randomly placed within each run the fixation cross changed to one of the letters A, T, or V followed by a question mark. This prompted participants to report if the most recent stimulus (directly before appearance of the letter) in the auditory (letter A), somatosensory (letter T for "tactile"), or visual (letter V) modality was presented with low or high intensity. The right foot was used to press either a left or a right pedal, and the pedal assignment (left = low/ right = high or left = high/right = low) was counterbalanced across participants.

It should be noted that our MMR paradigm in form of an attended roving stimulus sequence with relatively long ISI (1.15 s) differs from the classic oddball protocol for MMN elicitation in which participants are engaged in a primary task, often attending a separate modality.

Since this is not easily possible with our paradigm (containing auditory, somatosensory, and visual stimuli), we used catch-trials in each modality to ensure that attentional resources were distributed largely equally across the simultaneous stimulus streams. The ISI at the upper end of the range used for MMN elicitation was set during piloting such that the perceptually demanding tri-modal bilateral stimulation was deemed not to be overwhelming in terms of sensory overload.

2.3 | Probabilistic sequence generation

Each of the three sensory modalities (A, S, V) were presented as binary (low/high) stimulus sequences originating from a common probabilistic model. The model consists of a state s at time t evolving according to a Markov chain ($p(s_t|s_{t-1})$) with each state deterministically emitting a combination of three binary observations conditional on the preceding observation combination ($p(o_{A,t}, o_{S,t}, o_{V,t}|o_{A,t-1}, o_{S,t-1}, o_{V,t-1})$). For example, a transition expressed as [100|000] indicates a unimodal auditory change ($o_{A,t} = 1, o_{A,t-1} = 0$) with repeating somatosensory and visual modalities ($o_{S,t} = o_{S,t-1} = 0$ and $o_{V,t} = o_{V,t-1} = 0$). For each stimulus modality, in each state, the other two modalities form either congruent observations ([00] and [11]), or incongruent observations ([01] and [10]), which was used to manipulate the predictability of transitions in the sequences in different runs of the experiment.

Three types of stimulus sequences, depicted in Figure 2 were generated with different probability settings. The settings determine the transition probabilities within each modality given the arrangement of the other two modalities (i.e., either congruent or
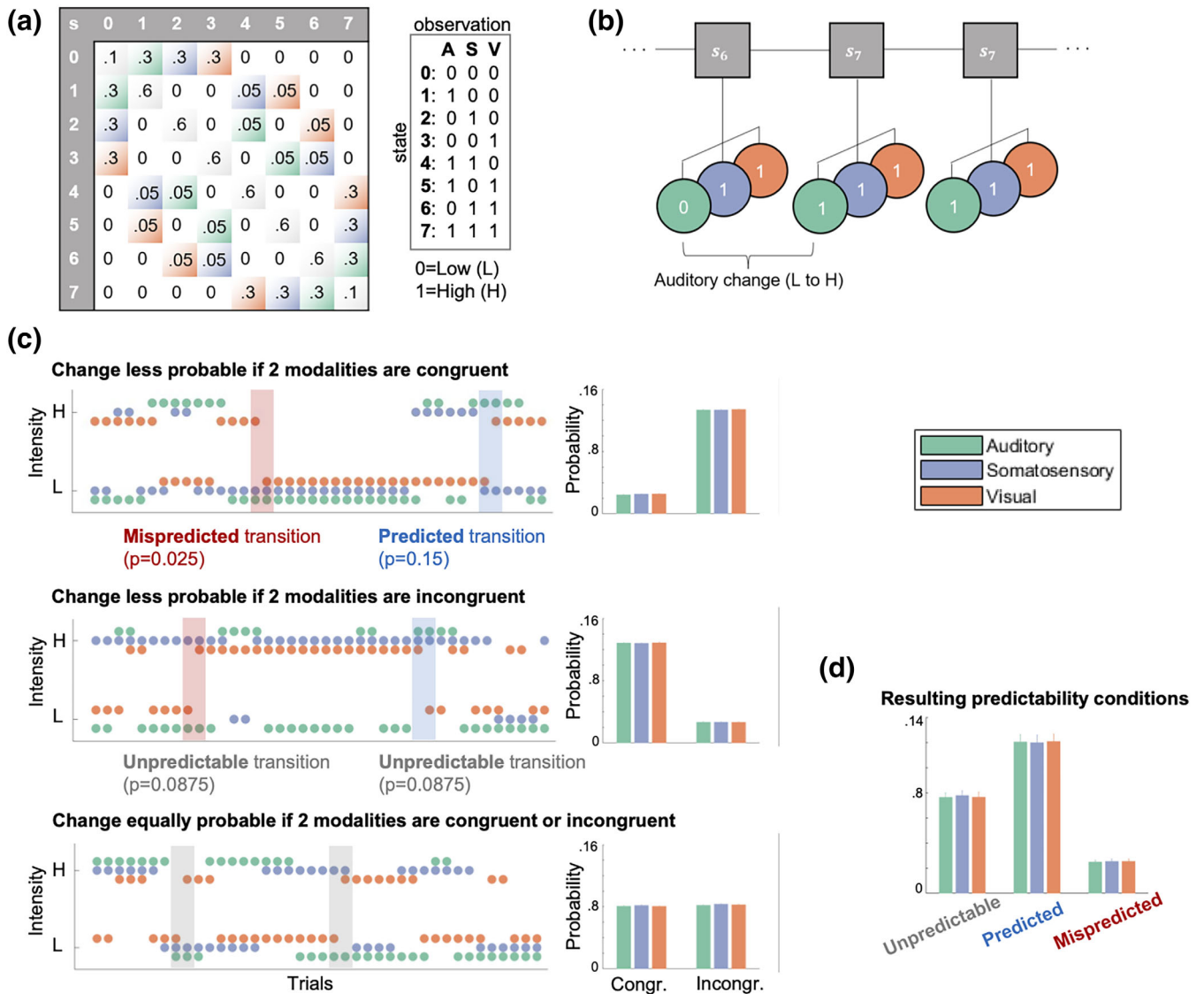
FIGURE 2 Probabilistic sequence generation. (a) Schematic of state transition matrix (left). Light color shading depicts transitions in the respective modality which were assigned specific transition probabilities: Green = auditory change, purple = somatosensory change, orange = visual change, gray diagonal = tri-modal repetition, white = multimodal change (set to zero). States 0–7 correspond to a specific stimulus combination (right), that is, eight permutations of low (0) and high (1) stimuli across three modalities (A = auditory; S = somatosensory; V = visual) as described in the main text. (b) Visualization of states (s) evolving according to a Markov chain emitting tri-modal binary outcomes. (c) Probability settings of stimulus sequences. Left column: Sequences. Right column: Averaged empirical change probabilities across all sequences. Top: Transition probabilities determine that for each modality a change is unlikely ($p = .025$) if the other two modalities are congruent (and likely if they are incongruent; $p = .15$). Middle: Transition probabilities determine that for each modality a change is likely ($p = .15$) if the other two modalities are congruent (and unlikely if they are incongruent; $p = .025$). Bottom: Transition probabilities determine that for each modality a change is equally likely ($p = .0875$) if the other two modalities are congruent or incongruent. (d) Averaged empirical change probabilities for predictability conditions.

incongruent). One setting defines lower change probability if the other two modalities are congruent (e.g., for any change in modality A from $t-1$ to t, S and V were *congruent* with $p(100|000) = p(000|100) = p(111|011) = p(011|111) = 0.025$ and S and V were *incongruent* with $p(101|001) = p(001|101) = p(110|010) = p(010|110) = 0.15$). The second setting defines lower change probability if the other two modalities are incongruent (e.g., for any change in modality A from $t-1$ to t, S and V were *incongruent* with $p(101|001) = p(001|101) = p(110|010) = p(010|110) = 0.025$ and S

and V were *congruent* with $p(100|000) = p(000|100) = p(111|011) = p(011|111) = 0.15$). The third setting defines equal change probability if the other two modalities are congruent or incongruent (e.g., for any change in modality A from $t-1$ to t, S and V were *congruent* with $p(100|000) = p(000|100) = p(111|011) = p(011|111) = 0.0875$ and S and V were *incongruent* with $p(101|001) = p(001|101) = p(110|010) = p(010|110) = 0.0875$).

In each of six experimental runs, the stimulus sequence was defined by one of the three different probability settings. Each

probability setting was used twice during the experiment and the order of the six different sequences was randomized. Participants were unaware of the sequence probabilities and any learning of sequence probabilities was considered to be implicit and task irrelevant.

Following the nomenclature suggested by Arnal and Giraud (2012), the resulting stimulus transitions for each modality within the different sequences can be defined as being either *predicted* (here higher change probability conditional on congruency/incongruency), *mispredicted* (here lower change probability conditional on congruency/incongruency) or *unpredictable* (here equal change probability). For each modality, repetitions are more likely ($p = .825$) than changes ($p = .175$) regardless of the type of probability setting and stimulus, resulting in classic roving standard sequences for each modality (mean stimulus train length: 5, mean range of train length: 2–34 stimuli).

2.4 | EEG data collection and preprocessing

Data were collected using a 64-channel active electrode EEG system (ActiveTwo, BioSemi, Amsterdam, Netherlands) at a sampling rate of 2048 Hz, with head electrodes placed in accordance with the extended 10–20 system. Individual electrode positions were recorded using an electrode positioning system (zebris Medical GmbH, Isny, Germany).

Preprocessing of the EEG data was performed using SPM12 (Wellcome Trust Centre for Neuroimaging, Institute for Neurology, University College London, London, UK) and in-house MATLAB scripts (MathWorks, Natick, MA). First, the data were referenced against the average reference, high-pass filtered (0.01 Hz), and down-sampled to 512 Hz. Subsequently, eye-blinks were corrected using a topographical confound approach (Berg & Scherg, 1994; Ille et al., 2002). The Data were epoched using a peri-stimulus time interval of −100 to 1050 ms and all trials were visually inspected and arti-factual data removed. Likewise, catch trials were omitted for all further analyses. Furthermore, the EEG data of two consecutive participants were found to contain excessive noise due to hardware issues, resulting in their exclusion from further analyses and leaving data of 32 participants. Finally, a low-pass filter was applied (45 Hz) and the preprocessed EEG data were baseline corrected with respect to the pre-stimulus interval of −100 to −5 ms. To use the general linear model (GLM) implementation of SPM, the electrode data of each participant were linearly interpolated into a 32×32 grid for each time point, resulting in one three-dimensional image (with dimensions $32 \times 32 \times 590$) per trial. These images were then spatially smoothed with a 12×12 mm full-width half-maximum Gaussian kernel to meet the requirements of random field theory, which the SPM software uses to control the family wise error rate.

2.5 | Event-related responses and statistical analysis

First, to extract basic MMR signals of each modality from the EEG data, we contrasted standard and deviant trials of each modality with

paired t-tests corrected for multiple comparisons by using cluster-based permutation tests implemented in fieldtrip (Maris & Oostenveld, 2007). Two time windows of interest were defined based on the literature (Duncan et al., 2009) to search for earlier negative clusters between 50 and 300 ms, corresponding to the MMN, and later positive clusters between 200 and 600 ms, corresponding to the P3. Clusters were defined as adjacent electrodes with a cluster defining threshold of $p_{fwe} < .05$.

For further analyses, GLMs were set up as implemented in SPM12, which allows defining conditions on the single trial level. To test for effects of stimulus repetitions on standards, deviants, and MMRs (deviants minus standards), a *TrainLength* model was defined that consisted of 45 regressors: an intercept regressor, 36 regressors coding for the repetition train length (trials binned into 1, 2, 3, 4–5, 6–8, >8 repetitions) for standards (i.e., the position of the standard in the current train) and deviants (i.e., the number of standards preceding the deviant) in each modality, as well as four global standard and four global deviant regressors. The binning was chosen such that trials were roughly balanced across the bins, corresponding to, on average, around 100 deviant trials and around 500 standard trials in each category for each modality The global regressors captured the train length (1, 2, 3, >3 repetitions) of standards and deviants regardless of their modality, meaning that trials in which standards occurred in all three modalities were coded as global standards, whereas trials in which a deviant occurred in any of the three modalities were coded as global deviants.

To test for the implicit effect of cross-modal predictability based on the different conditional probability setting in the sequence, a *Predictability* model was defined that consisted of 37 regressors: an intercept regressor and 18 regressors coding standards and deviants of each modality for each of the three conditions described above: *unpredictable* (trials originate from sequences with no conditional dependence between modalities), *predicted* (trials originate from sequences with conditional dependence; trials defined by change being likely), *mispredicted* (trials originate from sequences with conditional dependence; trials defined by change being unlikely). On the single-participant level, these were coded for congruent and incongruent trials separately resulting in 36 regressors. By definition, the number of trials in regressors with *mispredicted* trials was lowest, on average corresponding to around 60 deviant trials and around 800 standard trials per modality.

Finally, a *P3-Conjunction* model was specified that consisted of seven regressors: an intercept regressor and six regressors coding all standards and deviants for each of the three modalities. This model was used to apply SPM's second level conjunction analysis, contrasting standards and deviants across modalities in search of common P3 effects across modalities.

Each GLM was estimated on the single-trial data of each participant using restricted maximum likelihood estimation. This yielded β-parameter estimates for each model regressor over (scalp-) space and time, which were subsequently analyzed at the group level. Second level analyses consisted of a mass-univariate multiple regression analysis of the individual β scalp-time images with a design matrix specifying regressors for each condition of interest as well as a subject

factor. Second level beta estimates were contrasted for statistical inference and multiple comparison correction was achieved with SPM's random field theory-based FWE correction (Kilner et al., 2005).

2.6 | Source localization

To investigate the most likely underlying neuronal sources for the MMN and P3 MMR we applied distributed source reconstruction as implemented in SPM12 to the ERP data. For each participant, the MMN of each modality (auditory, somatosensory, visual) was source localized within a time window of 100–200 ms. For the P3, the average MMR at 330 ms was chosen for source localization as this time point showed the strongest overlap of P3 responses between modalities (based on the results of the P3 conjunction contrast).

Participant-specific forward models were created using an 8196-vertex template cortical mesh co-registered with the individual electrode positions via fiducial markers. An EEG head model based on the boundary element method was used to construct the forward model's lead field. For the participant-specific source estimates, multiple sparse priors under group constraints were applied. The source estimates were subsequently analyzed at the group level using the GLM implemented in SPM12. Second-level contrasts consisted of one-sample t tests for each modality as well as (global) conjunction contrasts across modalities. The resulting statistical parametric maps were thresholded at the peak level with $p < .05$ after FWE correction. The anatomical correspondence of the MNI coordinates of the cluster peaks were identified via cytoarchitectonic references using the SPM Anatomy toolbox (Eickhoff et al., 2005).

2.7 | Single-trial modeling of EEG data

In addition to the analysis of event-related potentials, the study aimed to compare different computational strategies of sequence processing potentially underlying neuronal generation of MMRs. To this end, we generated regressors from different BL models as well as a train length dependent change detection (TLCD) model making different predictions for the single-trial EEG data.

Theories on MMN generation hypothesize adaptation and memory-trace dependent change detection to contribute to the MMN. With prior repetition of stimuli, the response to standard stimuli tends to decrease while the response to deviant stimuli tends to increase. We defined the TLCD model to reflect such reciprocal dynamics of responses to stimulus repetition and change without invoking assumptions of probabilistic inference. The model is defined for each modality separately and tracks the stimulus train lengths c for a given modality by counting stimulus repetitions: $c_t = d_t(c_{t-1} + d_t)$, where $d_t = 1_{o_t = o_{t-1}}$ takes on the value 1 whenever the current observation o_t is a repetition of the previous observation o_{t-1} and $d_t = 0$ resets the current train length to zero. To form single-trial predictors of the EEG data, the model outputs values that increase linearly with train length and have opposite signs for standards and deviants:

$$\text{TLCD}(o_t) = \begin{cases} -ct & \text{if } d_t = 1 \\ c_{t-1} & \text{if } d_t = 0 \end{cases}$$

In addition to the TLCD model, different BL models were created to contrast the static train length based TLCD model with dynamic generative models tracking transition probabilities. The BL models consist of conjugate Dirichlet-Categorical models estimating probabilities of observations read out by three different surprise functions: Bayesian surprise (BS), predictive surprise (PS), and confidence-corrected surprise (CS).

BS quantifies the degree to which an observer adapts their generative model to incorporate new observations (Baldi & Itti, 2010; Itti & Baldi, 2009) and is defined as the Kullback–Leibler (KL) divergence between the belief distribution prior and posterior to the update: $BS(y_t) = KL\left(p(s_{t-1}|y_{t-1},...,y_1) \| p(s_t|y_t,...,y_1)\right)$. PS is based on (Shannon, 1948) definition of information and defined as the negative logarithm of the posterior predictive distribution, assigning high surprise to observed events y_t with low estimated probability of occurrence: $PS(y_t) = -\ln p(y_t|s_t) = -\ln p(y_t|y_{t-1},...,y_1)$. CS additionally considers the commitment of the generative model and scales with the negative entropy of the prior distribution (Faraji et al., 2018). It is defined as the KL divergence between the (informed) prior distribution at the current time step and a flat prior distribution $\widehat{p}(s_t)$ updated with the most recent event y_t: $CS(y_t) = KL(p(s_t)|\widehat{p}(s_t|y_t))$.

Following Faraji et al. (2018) surprise quantifications can be categorized as puzzlement or enlightenment surprise. While puzzlement refers to the mismatch between sensory input and internal model belief, closely related to the concept of prediction error, enlightenment refers to the update of beliefs to incorporate new sensory input. In the current study, we were interested in a quantification of the model inadequacy by means of an unsigned prediction error as reflected by surprise. As such, throughout the manuscript, with prediction error we do not refer to the specific term of (signed) reward prediction error as used for example in reinforcement learning but rather use it to refer to the signaling of prediction mismatch. While both PS and CS are instances of puzzlement surprise, CS is additionally scaled by belief commitment and quantifies the concept that a low-probability event is more surprising if commitment to the belief (of this estimate) is high. BS, on the other hand, is an instance of enlightenment surprise and is considered a measure of the update to the generative model resulting from new incoming observations.

A detailed description of the Bayesian observer, its transition probability version as well as the surprise read-out functions can be found in our previous work on somatosensory MMRs (Gijsen et al., 2021). Here, we will primarily provide a brief description of the specifics of two implementations of Dirichlet-Categorical observer models, a unimodal and a cross-modal model. Both observer models receive stimulus sequences (of one respective modality) as input and iteratively update a set of parameters with each new incoming observation. In each iteration, the estimated parameters are read out by the surprise functions (BS, PS, and CS) to produce an output which is subsequently used as a predictor for the EEG data.

For each modality, the unimodal Dirichlet-Categorical model considers a binary sequence with two possible stimulus identities (low and high) estimating transition probabilities with $y_t = o_t$ for $t = 1, ..., T$ with a set of hidden parameters $s^{(i)}$ for each possible transition from $o_{t-1} = i$. This unimodal model does not capture any cross-modal dependencies in the sequence (i.e., the alternation and repetition probabilities conditional on the tri-modal stimulus configuration). Therefore, we defined a cross-modal Dirichlet-Categorical model to address the question whether the conditional dependencies were used by the brain during sequence processing for prediction of stimulus change. The dependencies in the sequence were independent of stimulus identity but provide information about the probability of repetition or alternation (d_t) conditional on the congruency of the other modalities. The cross-modal model thus estimates alternation probabilities ($y_t = d_t$ for $t = 2, ..., T$) with a set of hidden parameters $s^{(i)}$ when other modalities are incongruent and $s^{(c)}$ when other modalities are congruent. Therefore, while the unimodal model learns the probability of stimulus *transitions* within modality, the cross-modal model learns the probability of stimulus *alternations* within modality conditional on the congruency of the other modalities. As such, the cross-modal model provides a minimal implementation of a Bayesian observer that captures the cross-modal dependencies in the sequences.

2.7.1 | Model fitting procedure

The technical details of the model fitting and subsequent Bayesian model selection (BMS) procedures are identical to Gijsen et al. (2021) where the interested reader is kindly referred to for further information. First, the stimulus sequence-specific regressor of each model was obtained for each participant. After z-score normalization, the regressors were fitted to the single-trial, event-related electrode data using a free-form variational inference algorithm for multiple linear regression (Flandin & Penny, 2007; Penny et al., 2003; Penny et al., 2005). The obtained model-evidence maps were subsequently subjected to the BMS procedure implemented in SPM12 (Stephan et al., 2009) to draw inferences across participants with well-established handling of the accuracy-complexity trade-off (Woolrich, 2012).

In total, eight regression models were fit: A null model (offset only), a TLCD regression model and, for each of the three surprise read-out functions, one regression model including only the unimodal regressors and one additionally including the cross-modal regressors. The purely unimodal regression model will be called UM and the regression model including unimodal and, additionally, cross-modal regressors will be called UCM. The design matrix of the TLCD regression model consisted of four regressors, an offset and the predicted parametric change responses for each of the three modality sequences (auditory, somatosensory, visual). Similarly, the design matrix of the UM regression model consisted of four regressors, an offset and the surprise responses of the unimodal model for each of the three modalities. The UCM regression model was identical to the UM regression model but with an additional three regressors

containing the cross-modal surprise responses for each modality. Therefore, the UCM regression model is more complex and gets assigned higher model evidence than the reduced UM regression model only if the additional regressors contribute significantly to a better model fit (Stephan et al., 2009).

To allow for the possibility of different timescales of stimulus integration (Maheu et al., 2019; Ossmy et al., 2013; Runyan et al., 2017), the integration parameter τ of the Dirichlet-Categorical model was optimised for each model, participant and peri-stimulus time-bin before model selection. To this end, model regressors were fit for a range of 11 tau parameter configurations ([0, 0.001, 0.0015, 0.002, 0.003, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2]) corresponding to integration windows with a 0.5 stimulus weighting at (half-life of) [600, 462, 346, 231, 138, 69, 34, 13, 6, 3] stimuli, of which the parameter with the best model evidence was chosen.

2.7.2 | Bayesian model comparison

The estimated model-evidence maps were used to evaluate the models' relative performance across participants via family-wise BMS (Penny et al., 2010). The model space was partitioned into three types of families to draw inference on different aspects of the involved models. Given that the literature provides some evidence for each of the three surprise read-out functions (BS, PS, CS) to capture some aspect of EEG MMRs, we included all of them in the family wise comparisons to avoid biasing the comparison of different BL models.

The first model comparison considered the full space of BL models as a single family (BL family) and compared it to the TLCD model (TLCD family) and the null model (NULL family). Since the BL models had their tau parameter optimized, which was not possible for the TLCD model, we applied the same penalization method used in our previous study (Gijsen et al., 2021). The degree to which the optimization on average inflated model evidence was subtracted from the BL models prior to BMS. Specifically, for all parameter values, the difference between the average model evidence and that of the optimized parameter was computed and averaged across poststimulus time bins, electrodes and participants.

Subsequent analyses grouped the different BL models into separate families: The second comparison grouped the BL models into two families of UM and UCM models, as well as the null model, to test which electrodes and time points showed influences of unimodal versus cross-modal processing. The third comparison grouped the BL models into three surprise families and the null model, to test whether the observed MMRs were best captured by PS, BS, or CS.

3 | RESULTS

3.1 | Behavioral results

Participants showed consistent performance in responding to the catch trials during each experimental run, indicating their ability to

globally maintain their attention to the tri-modal stimulus stream. Of the 85.5% responses made in time, 75.3% were correct with an average reaction time of 1.4 ± 0.25 s.

3.2 | Event-related potentials

3.2.1 | Unimodal MMRs

Cluster-based permutation tests confirmed the presence of early modality specific MMN components as well as later P3 MMRs for all three modalities. Both early and late MMRs showed a modulation by the number of stimulus repetitions, the details of which will be described in the following sections.

3.2.2 | Auditory MMRs

The MMN, as the classic MMR, has originally been studied in the auditory modality and is commonly described as the ERP difference wave calculated by subtraction of standard trials from deviant trials (deviants-standards). This difference wave typically shows a negative deflection at fronto-central electrodes and corresponding positivity at temporo-parietal sites, ranging from around 100 to 250 ms (Naatanen et al., 1978; Naatanen et al., 2007). Correspondingly, we find a significant negative fronto-central auditory MMN cluster between 80 and 200 ms (Figure 3a). Within the MMN cluster, deviants appear to deflect from the standard ERP around the peak of the auditory N1 component and reach their maximum difference around the peak of the subsequent P2 component. In the later time window, we observe positive MMRs at central electrodes between 200 and 400 ms, corresponding to a P3 modulation, as well as beyond 400 ms at progressively more posterior electrodes.

Within early and late auditory MMR clusters, the response to both standards and deviants was modulated by the number of standard repetitions. The auditory system is known to be sensitive to stimulus repetitions, particularly within the roving standard paradigm (Baldeweg et al., 2004; Cowan et al., 1993; Ulanovsky et al., 2003; Ulanovsky et al., 2004). Therefore, we hypothesized a gradual increase of the auditory response to standard stimuli around the time of the MMN, known as repetition positivity (Baldeweg, 2006; Baldeweg et al., 2004; Haenschel et al., 2005) as well as reciprocal negative modulation of the corresponding deviant response (Bendixen et al., 2007; Naatanen et al., 2007). Together, these effects should result in a gradual increase of the MMN amplitude with stimulus repetition. Indeed, linear contrasts applied to the GLM beta parameter estimates of the *TrainLength* model revealed that the MMN increases with the repetition of standards before a deviant was presented (94–200 ms, cluster $p_{fwe} < .001$). This effect was driven by a negative linear modulation of the deviant response (98–200 ms, cluster $p_{fwe} < .001$) as well as a repetition positivity effect on the standards (111–172 ms, cluster $p_{fwe} < .001$). Similarly, the later P3 MMR increased with standard repetitions (200–600 ms, cluster $p_{fwe} < .001$) and this effect was driven by an increase of deviant responses (200–600 ms, cluster $p_{fwe} < .001$) and a decrease of standard responses (205–359 ms, cluster $p_{fwe} < .001$). Given the temporal difference between standard (around 200–350 ms) and deviant (200–600 ms) train length effects, the parametric modulation of the late MMR beyond 350 ms seems to be primarily driven by the increase in deviant responses.

3.2.3 | Somatosensory MMRs

We hypothesized somatosensory MMRs to consist of early bilateral (fronto-) temporal negativities, resulting primarily from increased N140 components (Kekoni et al., 1997), with a corresponding central positivity extending into a later central P3 component.

After an early mismatch effect starting at ~50 ms at fronto-central electrodes, a more pronounced bilateral temporal cluster emerged that extended from ~90 to 190 ms and can be considered the somatosensory equivalent of the auditory MMN (Figure 3b). A reversed positive central component can be observed at the time of the somatosensory MMN (sMMN) and throughout the entire later time window (200–600 ms) at which point it can be considered a putative P3 MMR.

Early and late somatosensory MMRs were significantly modulated by stimulus repetition. Bilateral electrodes within the sMMN cluster show an increase of the sMMN amplitude with repetition (123–166 ms, cluster $p_{fwe} < .05$). This effect was driven by an increase of deviant negativity (135–188 ms, cluster $p_{fwe} < .05$) in combination with a positivization of the standard (86–188 ms, cluster $p_{fwe} < .05$). Similarly, the later P3 MMR increases with repetition of standards (200–600 ms, cluster $p_{fwe} < .05$), mutually driven by increasing deviant responses (221–600 ms, cluster $p_{fwe} < .05$) and decreasing standard responses (200–600 ms, cluster $p_{fwe} < .05$).

3.2.4 | Visual MMRs

We hypothesized visual MMRs to present as an early MMN at occipital to parieto-temporal electrodes and a later P3 component at central electrodes. Although less pronounced than its auditory and somatosensory counterparts, we indeed observed a negative visual mismatch component that developed from occipital to parieto-temporal electrodes between ~130 and 200 ms (Figure 3c). In the later time window, we found a central positive component between ~300 and 600 ms, corresponding to a P3 MMR.

Within the significant visual MMN (vMMN) cluster, the linear contrast testing for repetition effects did not reach significance when correcting clusters for multiple comparisons ($p_{fwe} > .05$). However, it is worth noting that some electrodes in this cluster seemed to show a similar pattern of response increases and decreases as in the auditory and somatosensory modality, which became apparent at more lenient thresholds. The vMMN tended to become more negative with repetition of standards (143–148 ms, peak $p_{uncorr} < .005$), with opposite tendencies of deviant negative increase (143–148 ms, peak $p_{uncorr} < .05$)
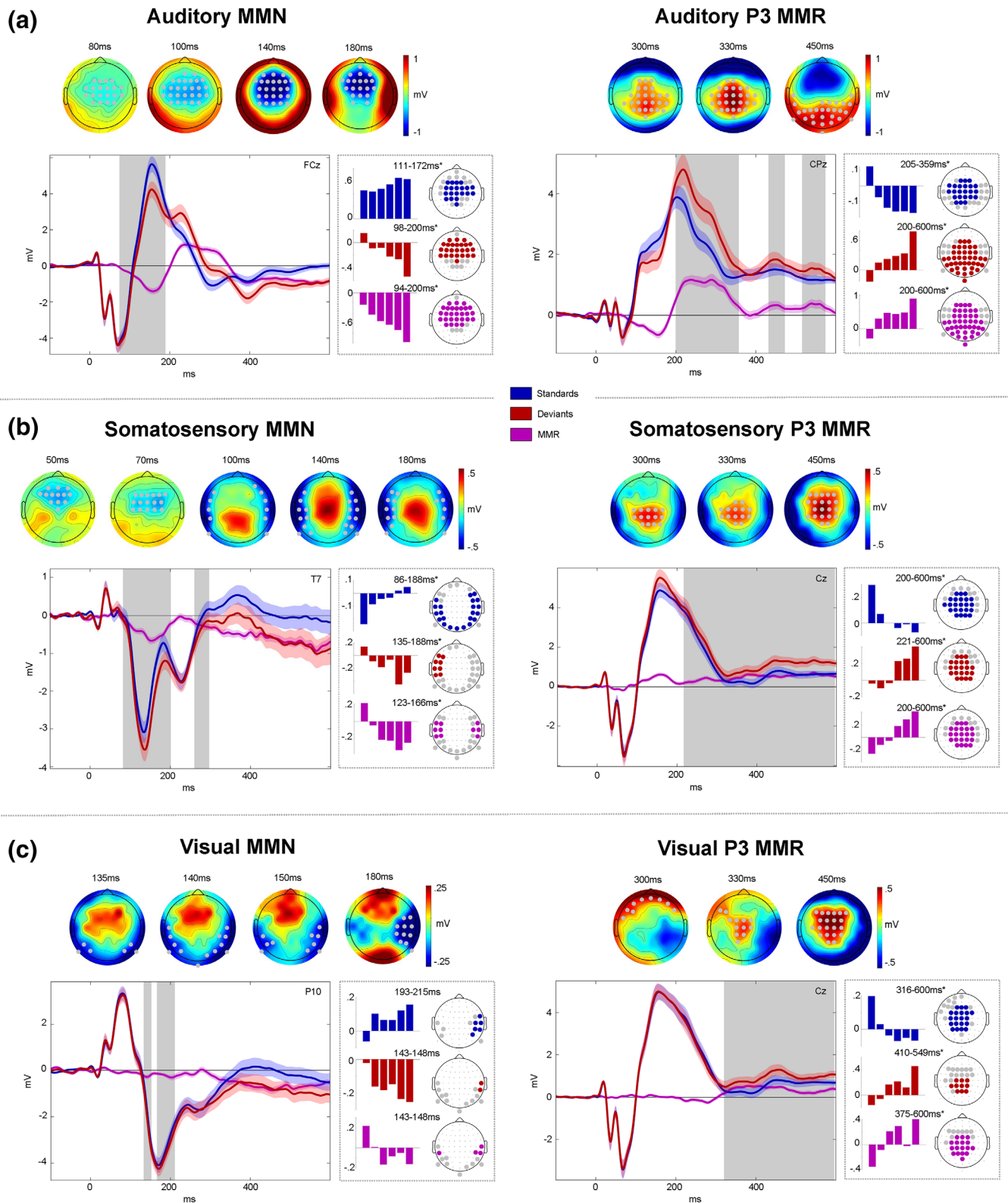
FIGURE 3 Mismatch responses. Panels (a–c) show mismatch responses (MMRs) of auditory (a), somatosensory (b), and visual (c) modalities. Within panels: Left: mismatch negativity (MMN). Right: P3 MMR. Gray dots (top) and gray boxes (bottom) indicate significant MMR electrodes and time points with $p_{fwe} < .05$. Top row: MMR scalp topographies (deviants-standards). Bottom row: Grand average ERPs (left panels) and beta parameter estimates of significant linear contrast clusters (right panels). Colored bars depict six beta parameter estimates of the *TrainLength* GLM (1, 2, 3, 4–5, 6–8, >8 repetitions) averaged across electrodes within linear contrast clusters. Asterisks indicate significance of the linear contrast ($p_{fwe} < .05$).
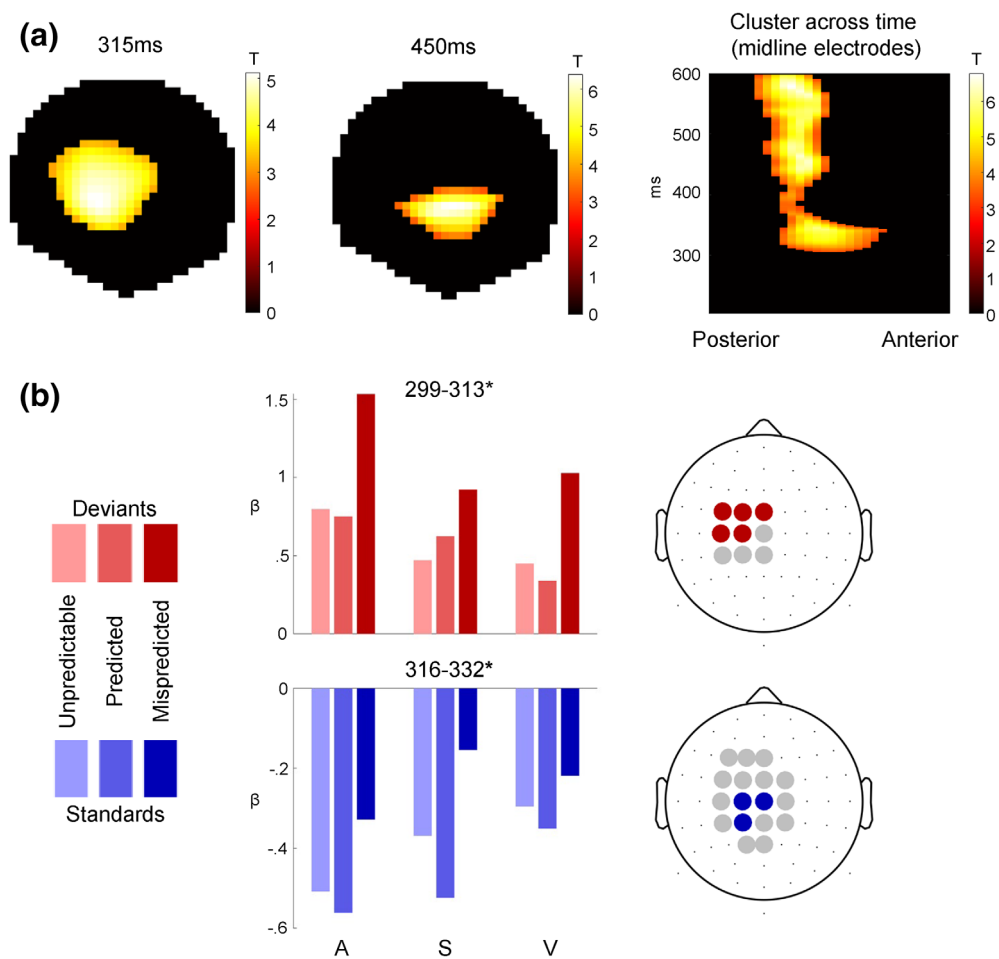
FIGURE 4 Cross-modal P3 effects. (a) T-Maps of the conjunction of deviant > standard contrasts across the auditory, somatosensory, and visual modalities. (b) Beta estimates averaged across electrodes within significant clusters with peak p_{fwe}<0.05, resulting from two-way ANOVAs testing for differences between *unpredictable*, *predicted*, and *mispredicted* deviants (red) and standards (blue).

and standard decrease (193–215 ms, peak p_{fwe} < .05). Thus, although we cannot conclude a modulation by standard repetition of the vMMN with any certainty, the observed beta parameters are in principle compatible with the effects observed in the auditory and somatosensory modalities (please see the discussion for potential reasons for the reduced vMMN in our data).

Within the P3 MMR cluster, on the other hand, we find significant clusters of linear increase of the MMR (375–600 ms, cluster p_{fwe} < .05), again constituted by an increase in deviant responses (410–549 ms, cluster p_{fwe} < .05) and concomitant decrease in standard responses (316–600 ms, cluster p_{fwe} < .05).

3.2.5 | Cross-modal P3 effects

In search of a common P3 effect to deviant stimuli, we created conjunctions of the *deviants > standards* contrasts across the auditory, somatosensory, and visual modalities. The conjunction revealed a common significant cluster starting at ~300 ms (cluster p_{fwe} < .05) that comprised anterior central effects around 300–350 ms followed by more posterior effects from 400 to 600 ms (Figure 4a).

To investigate the modulation of the P3 MMR by predictability, we used two-way ANOVAs with the three-level factor *modality*

(*auditory*, *somatosensory*, *visual*) and the three-level factor *predictability* condition (*predicted*, *mispredicted*, *unpredictable*). Separate ANOVAs were applied to deviants and standards. We hypothesized that the cross-modal P3 MMR might be sensitive to multisensory predictive information in the sequence, as the P3 has been shown to be sensitive to global sequence statistics (Bekinschtein et al., 2009; Wacongne et al., 2011) and to be modulated by stimulus predictability (Horvath et al., 2008; Horvath & Bendixen, 2012; Max et al., 2015; Prete et al., 2022; Ritter et al., 1999; Sussman et al., 2003). Indeed, within the common P3 cluster, both deviants (299–313 ms, peak p_{fwe} < .05) and standards (316–332 ms, peak p_{fwe} < .05) show significant differences between predictability conditions. No significant interaction of predictability condition with modality was observed.

Post hoc *t* tests were applied to the peak beta estimates to investigate the differences between the three pairs of conditions. For the ANOVA concerning the deviant trials, post hoc *t* tests show a significant difference for *mispredicted > predicted* (*t* = 14.667; *p* < .001, Bonferroni corrected), *mispredicted > unpredictable* (*t* = 14.76; *p* < .001, Bonferroni corrected) and no significant difference between *unpredictable > predicted* conditions (*t* = 0.01; *p* > .05). Similarly, for the ANOVA concerning the standard trials, post hoc *t* tests show that there is a significant difference for *mispredicted > predicted* (*t* = 10.67;

FIGURE 5 Source localization. Top row: significant sources (p_{fwe} < .05) for the auditory (green), somatosensory (purple), and visual (orange) mismatch negativity (MMN) as well as their conjunction (yellow). Bottom row: significant sources (p_{fwe} < .05) for the conjunction (yellow) of the P3 mismatch response (MMR) in the auditory, somatosensory, and visual modalities.
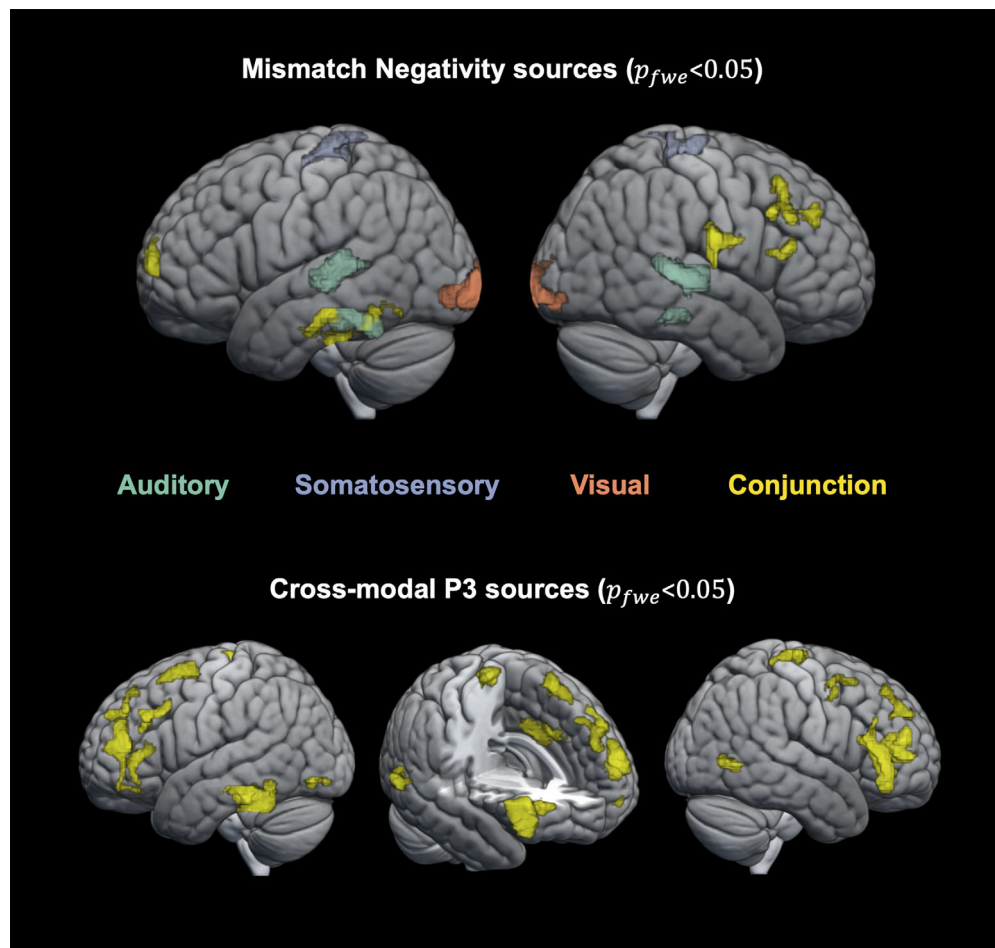
p < .001, Bonferroni corrected), *mispredicted > unpredictable* ($t = 6.87$; p < .001, Bonferroni corrected) and *unpredictable > predicted* conditions ($t = 3.83$; p < .001, Bonferroni corrected).

Taken together, this result suggests that stimuli which were mispredicted based on the predictive multisensory configuration resulted in increased responses within the common P3 cluster compared to predicted or unpredictable stimuli, regardless of their role as standards or deviants in the current stimulus train.

For completeness, we also tested the effect of predictability in the earlier MMN cluster, but we did not observe any significant modulations here (results not shown).

3.3 | Source localization

The source reconstruction analysis resulted in significant clusters of activation for each modality's MMN as well as the P3 MMR. The results are depicted in Figure 5 and cytoarchitectonic references are described in Table 1.

For each modality, the MMN was localized to source activations in the respective modality's sensory cortex and frontal cortex. Source localization of the auditory MMN shows the strongest activation in bilateral superior temporal areas (p_{fwe} < .05; left cluster: peak $t = 6.20$;

right cluster: peak $t = 7.64$) corresponding to auditory cortex and in inferior temporal areas (p_{fwe} < .05; left cluster: peak $t = 5.63$; right cluster: peak $t = 5.60$). The somatosensory MMN shows highest source activation in postcentral gyrus (p_{fwe} < .05; left cluster: peak $t = 5.22$; right cluster: peak $t = 4.92$) corresponding to primary somatosensory cortex. Similarly, the vMMN shows highest source activation in the occipital cortex (p_{fwe} < .05; left cluster: peak $t = 6.18$; right cluster: peak $t = 5.17$), around the occipital pole, corresponding to visual areas (V1–V4). Lowering the threshold to p_{uncorr} < .001 (only shown in Table 1) suggests additional activation of hierarchically higher sensory areas such as secondary somatosensory cortex for the sMMN (p_{uncorr} < .001; left cluster: peak $t = 4.21$; right cluster: peak $t = 5.01$) and lateral occipital cortex (fusiform gyrus) for vMMN (part of the primary visual cluster). In addition to the sensory regions, common frontal sources with dominance on the right hemisphere were identified using a conjunction analysis for the MMN of all three modalities. In particular, significant common source activations were found in the right inferior frontal gyrus (IFG; p_{fwe} < .05; cluster: peak $t = 3.15$) and right middle frontal gyrus (MFG; p_{fwe} < .05; cluster: peak $t = 2.89$). Additional significant common sources include frontal pole (p_{fwe} < .05; left cluster: peak $t = 2.56$; right cluster: peak $t = 2.28$), left inferior temporal gyrus (p_{fwe} < .05; cluster: peak $t = 2.52$) and right inferior parietal lobe (p_{fwe} < .05; cluster: peak $t = 2.85$).

TABLE 1 Source localization and cytoarchitectonic reference

| Contrast	Hemisphere	Cytoarchitecture (probability)	MNI coord. at cytoarch.	t-Statistics at cytoarch. (p-value)
aMMN	Left	Auditory areas:		
		TE 4 (67.3%)	−52 −26 0	6.2 (p_{fwe} < .05)
		TE 3 (15.1%)	−59.8 −17.6 5.4	4.81 (p_{fwe} < .05)
		TE 1 (50.7%)	−50.3 −19.2 5.8	4.05 (p_{uncorr} < .001)
aMMN	Right	Auditory areas:		
		TE 4 (54.1%)	−56 −26 0	7.64 (p_{fwe} < .05)
		TE 3 (33.6%)	−64.2 −16.4 5	5.85 (p_{fwe} < .05)
		TE 1 (61.9%)	53 −10.1 3.8	3.53 (p_{uncorr} < .001)
sMMN	Left	Somatosensory areas:		
		3b [S1] (31.4%)	−15.7 −33.7 68.6	4.8 (p_{fwe} < .05)
		OP4 [S2] (38.6%)	−65 −14.8 20.1	3.72 (p_{uncorr} < .001)
		OP1 [S2] (16%)		
sMMN	Right	Somatosensory areas:		
		3b [S1] (40.3%)	13.3 −33.7 68.2	4.76 (p_{fwe} < .05)
		OP4 [S2] (prob. 46.9%)	66.5 −10.6 20.9	3.51 (p_{uncorr} < .001)
		OP1 [S2] (prob. 9.1%)		
vMMN	Left	Visual areas:		
		hOc1 [V1] (84.6%)	−10 −100 0	6.18 (p_{fwe} < .05)
		hOc2 [V2] (11.1%)		
		hOc4v [V4] (51%)	−30 −87.9 -11.7	5.51 (p_{fwe} < .05)
		hOc3v [V3] (24.6%)		
		FG4 (89.5%)	−43.5 −49.4 −13.3	3.83 (p_{uncorr} < .001)
vMMN	Right	Visual areas:		
		hOc1 [V1] (86.4%)	20 −100 −4	5.17 (p_{fwe} < .05)
		hOc2 [V2] (11%)		
		hOc3 [V3] (43.7%)	34.4 −88.7 −7.7	4.7 (p_{fwe} < .05)
		hOc4 [V4] (23.8%)		
		FG4 (59.8%)	50.3 −42.5 −18.8	3.82 (p_{uncorr} < .001)
MMN conjunction	Left	Frontal pole (28%)	−14 62 8	2.56 (p_{fwe} < .05)
		Inferior temporal gyrus (49%)	−50 −24 −28	2.52 (p_{fwe} < .05)
MMN conjunction	Right	Middle frontal gyrus (36%)	42 28 30	2.89 (p_{fwe} < .05)
		Inferior frontal gyrus (53%)	53 27.6 17.3	2.96 (p_{fwe} < .05)
		Frontal pole (65%)	30 42 32	2.28 (p_{fwe} < .05)
		Inferior parietal lobe (46.8%)	62 −14 24	2.85 (p_{fwe} < .05)
P3 con-junction	(Left)	Anterior cingulate gyrus (51%)	−6 14 36	4.34 (p_{fwe} < .05)
P3 conjunction	Left	Frontal pole (74%)	−34 44 20	3.31 (p_{fwe} < .05)
		Inferior frontal gyrus (35%)	52 26 18	2.74 (p_{fwe} < .05)
		Middle frontal gyrus (78%)	−44 30 32	2.87 (p_{fwe} < .05)
		Superior frontal gyrus (43%)	−22 12 60	3.13 (p_{fwe} < .05)
		Inferior temporal gyrus (34%)	−48 −44 −24	3.21 (p_{fwe} < .05)
		Lateral occipital [hOc4la] (81%)	−46 −80 −8	2.99 (p_{fwe} < .05)
P3 conjunction	Right	Frontal pole (86%)	45 42 10	3.57 (p_{fwe} < .05)
		Inferior frontal gyrus (39.5%)	52 26 18	3.45 (p_{fwe} < .05)
		Middle frontal gyrus (38%)	36 2 54	3.0 (p_{fwe} < .05)
		Precentral gyrus [4a] (19%)	6 −32 64	2.64 (p_{fwe} < .05)
		Lateral occipital cortex [hOc5] (55%)	56 −62 0	3.45 (p_{fwe} < .05)
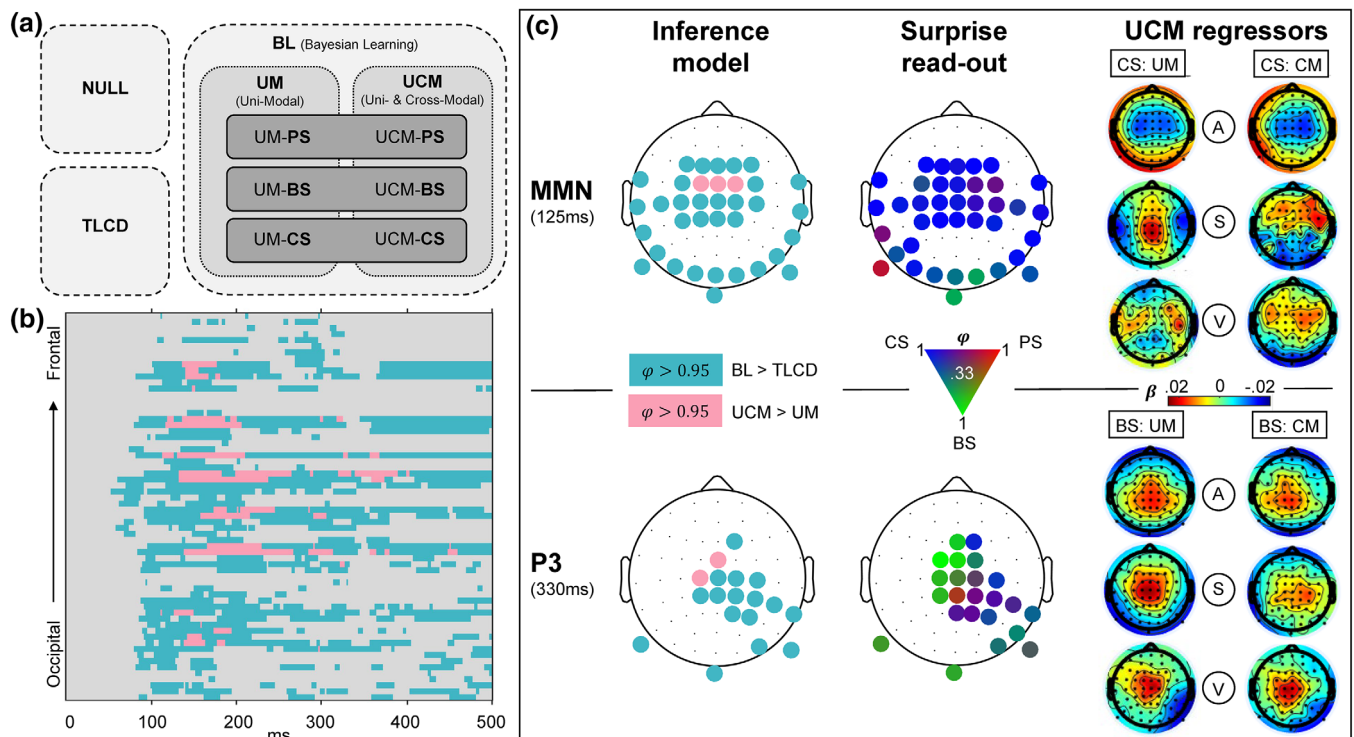
FIGURE 6 Modeling results. (a) Schematic overview of models. Model comparison 1 (light-gray box, dashed contour): Null model family (NULL), train length-dependent change detection model family (TLCD), and Bayesian learning model family (BL). Comparison 2 (gray box, dotted contour): Unimodal regression model family (UM), cross-modal regression model family (UCM). Comparison 3 (dark-gray box, line contour): Read-out model family comparison of predictive surprise family (PS), Bayesian surprise family (BS), and confidence-corrected surprise family (CS). (b) Results of comparison 1 and 2 shown for all electrodes and poststimulus time points. Color depicts exceedance probability (EP) $\varphi > 0.95$. Light-blue = BL > TLCD, pink = UCM > UM. (c) Topography of modeling results at time windows of MMN (top row) and P3 (bottom row). Left column: Results of comparison 1 (same colors as (B), depicting $\varphi > 0.95$). Middle column: Results of comparison 3. EPs between 0.33 and 1 of the three surprise functions are represented by a continuous three-dimensional RGB scale (red = predictive surprise [PS]; green = Bayesian surprise [BS]; blue = confidence-corrected surprise [CS]). Right column: Beta estimates of the model regressors of the UCM model (regressors: A = auditory; S = somatosensory; V = visual; CM = cross-modal; UM = unimodal) for CS read-out models (top) and BS read-out models (bottom).

For the late P3 MMR, a wide range of sources was expected to contribute to the EEG signal (Linden, 2005; Sabeti et al., 2016). To identify those that underlie the P3 MMR common to all modalities, we used a conjunction analysis. Significant clusters were found primarily in anterior cingulate cortex ($p_{fwe} < .05$; cluster: peak $t = 4.34$) and bilateral (pre-)frontal cortex ($p_{fwe} < .05$; left IFG cluster: peak $t = 3.57$; left superior frontal gyrus cluster: peak $t = 3.13$; left MFG cluster: peak $t = 2.87$; left frontal pole cluster: peak $t = 3.31$; right IFG cluster: peak $t = 3.45$; right MFG cluster: peak $t = 3.0$; right frontal pole cluster: peak $t = 3.57$). Additional significant sources were found in left inferior temporal gyrus ($p_{fwe} < .05$; cluster: peak $t = 3.21$), left and right lateral occipital cortex ($p_{fwe} < .05$; left cluster: peak $t = 2.99$; right cluster: peak $t = 3.45$) and right precentral gyrus ($p_{fwe} < .05$; cluster: peak $t = 2.64$).

3.4 | Single-trial modeling

As described in the previous sections, the responses of standards and deviants show specific sensitivity to (1) stimulus repetition and

(2) cross-modal conditional probability. To investigate the computational principles underlying these response profiles, eight different models capturing various learning strategies were fit to the single-trial EEG data and compared via family-wise BMS. A summary of the modeling results is depicted in Figure 6.

The first model comparison aimed to further investigate observation (1) and the question whether the observed parametric modulation of standard and deviant EEG responses merely reflects a combination of neuronal adaptation and change detection dynamics or if the observed response patterns are indicative of an underlying generative model engaged in probabilistic inference. To this end, we ran family wise BMS which is schematically depicted in Figure 6a. The first comparison concerned the TLCD model, a model family containing all BL models, and a null model. In the fronto-central, temporal, occipital, and central electrodes showing MMN and P3 effects, the model comparison shows strong evidence in favor of the BL model family with an exceedance probability $\varphi > 0.95$ (corresponding to expected posterior probability $\langle r \rangle > .7$) from ~70 ms onward. On the other hand, the TLCD model did not exceed $\varphi > 0.95$ for any electrode or time point. Therefore, the TLCD model was disregarded at this

point, and we focused further investigation on the different BL models.

The second comparison set out to investigate observation (2) and evaluate the contribution of stimulus alternation tracking conditional on multimodal configurations beyond unimodal transition probability inference. Within the electrodes and time-points with sufficient evidence for BL signatures (as established by the first model comparison), a comparison of a purely unimodal model family with a cross-modally informed model family (UCM) was performed. Those electrodes and time-points where the additional inclusion of cross-modal regressors (UCM models) provided better model fits than purely unimodal models are highlighted in Figure 6b,c. The UCM family outperforms the UM family at central and fronto-central electrodes at ∼100–400 ms (with $\varphi > 0.95$, corresponding to $\langle r \rangle > .7$).

Inspection of the beta estimates of auditory, somatosensory, and visual regressors of the UCM regression models shows that the beta maps of the unimodal predictors of the model resemble the ERP mismatch topographies of the respective modalities (depicted in Figure 6c). The cross-modal predictor, on the other hand, rather shows (fronto-) central activations which appear to resemble frontal aspects of the respective auditory, somatosensory, and visual MMRs.

The third comparison concerned the three surprise measures used as read-out functions for the probabilistic models. Overall, the family comparison does not show overwhelming evidence for any specific surprise function as only few electrodes reach exceedance probabilities of $\varphi > 0.95$. Nevertheless, a tendency of the MMN and the P3 to reflect different surprise dynamics can be observed. Although around the time of the MMN, only some electrodes show $\varphi > 0.95$ in favour of CS, inspection of the topographies without φ thresholding (as depicted in Figure 6c) shows CS to be dominant throughout the spatio-temporal range of the MMN (as suggested by higher EPs compared to BS and PS). On the other hand, at the time of the P3, BS appears to be the dominant surprise computation with multiple (fronto-) central electrodes showing $\varphi > 0.95$. Overall, the surprise comparison provides some evidence for a reflection of CS dynamics in the earlier mismatch signals around the time of the MMNs and suggests a tendency of the P3 to reflect BS dynamics.

In a final analysis, the optimal observation integration parameter τ was inspected. For each modality, the significant MMN clusters of the ERP analyses were used to inspect the optimal integration window of the regression models. For the UM regression model, highly similar optimal integration parameters were found within the electrodes and time-points of the different MMN clusters with no significant difference between the modalities. The optimal integration parameters were found to correspond to windows of stimulus integration with a half-life of (50% weighting at) around 5 to 25 stimuli (mean±SD τ across participants within the MMN clusters: UM-CS = 0.0608±0.0189 [8–16 stimuli], UM-BS = 0.0887±0.0184 [6–9 stimuli], UCM-CS = 0.0550± 0.0179 [9–18 stimuli], UCM-BS = 0.0954±0.0194 [6–9 stimuli]; mean±SD τ across participants within the P3 clusters: UM-CS = 0.0536±0.0214 [9–21 stimuli], UM-BS = 0.0811± 0.0191 [6–11 stimuli], UCM-CS = 0.0469±0.0196 [10–25 stimuli], UCM-BS = 0.0862±0.0215 [6–10 stimuli]). Overall, the same range of stimulus integration was found

for the UM and UCM regression models and CS models tended to have higher integration windows (∼10–20 stimuli) compared to Bayesian surprise models (∼5–10 stimuli).

4 | DISCUSSION

The present study set out to compare mismatch signals in response to tri-modal sequence processing in the auditory, somatosensory, and visual modalities and to investigate influences of predictive cross-modal information. We found comparable but modality specific signatures of MMN-like early mismatch processing between 100 and 200 ms in all three modalities, which were source localized to their respective sensory specific cortices and shared right lateralized frontal sources. An additional cross-modal signature of mismatch processing was found in the P3 MMR for which a common network with frontal dominance was identified. With exception of the vMMN, both mismatch signals (MMN and P3) show parametric modulation by stimulus train length driven by reciprocal tendencies of standards and deviants across modalities. Strikingly, standard and deviant responses within the cross-modal P3 cluster were sensitive to predictive information carried by the tri-modal stimulus configuration. Comparisons of computational models indicated that BL models, tracking transitions between observations, captured the observed dynamics of single-trial responses to the roving stimulus sequences better than a static model reflecting TLCD. Moreover, a BL model which additionally captured cross-modal conditional dependence of stimulus alternation outperformed a purely unimodal BL model primarily at central electrodes. The comparison of different read-out functions for the BL models provides tentative evidence that the early MMN may reflect dynamics of CS whereas later P3 MMRs seem to reflect dynamics of BS.

4.1 | Modality specific mismatch signatures in response to tri-modal roving stimuli

By using a novel tri-modal roving stimulus sequence originating from an underlying Markov process of state transitions, we were able to elicit and extract unique EEG signatures in each of the three sensory modalities (auditory, somatosensory, and visual).

Of the EEG mismatch signatures, the auditory MMN is one of the most widely researched responses to deviation from an established stimulus regularity (Naatanen et al., 1978; Winkler et al., 2009). Contrasting responses to standard and deviant stimuli of the auditory sequence in the current study resulted in the expected fronto-central MMN signature with more negative responses to deviants compared to standards. The extent of the MMN might suggest an underlying negative mismatch component as proposed by Naatanen et al. (2005), which drives a more negative going ERP around the N1, extending beyond the P2 component. Such post-N1 effects of the MMN have been suggested as markers of a "genuine" mismatch component in contrast to confounds by stimulus properties modulating auditory

ERP components (Naatanen et al., 2007) and might speak against pure N1 adaptation (as suggested by Jaaskelainen et al., 2004; May et al., 1999).

The somatosensory equivalent to the auditory MMN (sMMN) reported in the current study shows negative polarity at bilateral temporal electrodes and corresponding central positivity. The sMMN likely reflects an enhanced N140 component, as suggested by Kekoni et al. (1997). However, most previous sMMN studies used oddball paradigms where some critical discussion revolves around the distinction of the sMMN from an N140 modulation by stimulus properties alone. Here, we report an sMMN around the N140 which can be assumed to be independent of stimulus confounds due to the reversed roles of standard and deviant stimuli in the roving paradigm. Although several previous studies have reported somatosensory mismatch responses, conflicting evidence exists regarding the exact components that may constitute an equivalent to the auditory MMN. Some studies report a more fronto-centrally oriented negativity (Kekoni et al., 1997; Shen et al., 2018; Spackman et al., 2007; Spackman et al., 2010) or observed such pronounced central positivity that they were led to conclude that it is in fact the central positivity that should be considered the somatosensory equivalent of the aMMN (Akatsuka et al., 2005; Shinozaki et al., 1998). However, some evidence appears to converge on a temporally centered negativity with corresponding central positivity as the primary sMMN around 140 ms (Gijsen et al., 2021; Ostwald et al., 2012).

While the auditory and somatosensory MMN's in the current study were found to be highly comparable in their signal strength, their hypothesized counterpart in the visual modality showed a comparatively weaker response. Nevertheless, we found a significant vMMN at occipital electrodes extending to temporal electrodes within a time window of 100–200 ms poststimulus, with corresponding (fronto-) central positivity. This observation is in line with previous research reporting posterior (Cleary et al., 2013; Kimura et al., 2010; Urakawa et al., 2010) and temporal (Heslenfeld, 2003; Kuldkepp et al., 2013) patterns of vMMN with corresponding central positivity (Cleary et al., 2013; Czigler et al., 2006; File et al., 2017).

4.2 | Neuronal generators of MMN signatures

Source reconstruction analyses were used to identify underlying neuronal generators of the modality specific MMN signatures. Interestingly, for each sensory modality, we found generators in the primary and higher order sensory cortices as well as additional frontal generators in IFG and MFG.

The sensory specific neuronal sources underlying the auditory MMN were identified as bilateral auditory cortex with a dominance in hierarchically higher auditory areas. With an additional modality independent contribution of right lateralized frontal sources, this set of neuronal generators identified for the aMMN is in line with previous research suggesting primary auditory cortex and higher auditory areas in superior temporal sulcus as well as right IFG as underlying the aMMN (Garrido et al., 2008; Garrido, Kilner, Kiebel, & Friston, 2009;

Molholm et al., 2005; Naatanen et al., 2005; Opitz et al., 2002) with consideration of an additional frontal generator in MFG (Deouell, 2007).

The sources underlying the sMMN were identified in the current study as primary (S1) and secondary (S2) somatosensory cortices with additional frontal generators in right IFG and MFG. This finding is in accordance with previous research showing a combined response of S1 and S2 to underlie the sMMN (Akatsuka, Wasaka, Nakata, Kida, Hoshiyama, et al., 2007; Akatsuka, Wasaka, Nakata, Kida, & Kakigi, 2007; Andersen & Lundqvist, 2019; Butler et al., 2012; Gijsen et al., 2021; Naeije et al., 2016, 2018; Ostwald et al., 2012; Spackman et al., 2010) in combination with involvement of (inferior) frontal regions (Allen et al., 2016; Downar et al., 2000; Fardo et al., 2017; Huang et al., 2005; Ostwald et al., 2012).

For the visual modality, we identified sources in visual areas (V1–V4) and additional frontal activations in IFG and MFG as the neuronal generators underlying the vMMN. Previous studies have shown similar combinations of visual and prefrontal areas (Kimura et al., 2010; Kimura et al., 2011; Kimura et al., 2012; Urakawa et al., 2010; Yucel et al., 2007) and have particularly highlighted the IFG as a frontal generator of the vMMN (Downar et al., 2000; Hedge et al., 2015). Similarly, an fMRI study of perceptual sequence learning in the visual system has shown right lateralized prefrontal activation in addition to activations in visual cortex in response to regularity violations (Huettel et al., 2002). Yet another study has suggested a role for right prefrontal areas in interaction with hierarchically lower visual areas for the prediction of visual events (Kimura et al., 2012), all in line with our results.

Overall, our finding of inferior and middle frontal sources for the MMN in all three modalities provides further evidence for a modality independent role for these generators as previously suggested by Downar et al. (2000). As such, these modality-independent frontal generators might reflect higher stages of a predictive hierarchy working across modalities in interaction with lower modality specific regions, as previously suggested primarily for the auditory modality (Garrido, Kilner, Stephan, & Friston, 2009).

4.3 | Modulation of the MMN by stimulus repetition

An important feature of the MMN which theories of its generation have aimed to account for is its sensitivity to stimulus repetition. The MMN is known to increase with prior repetition of standards (Imada et al., 1993; Javitt et al., 1998; Naatanen & Näätänen, 1992; Sams et al., 1983). Correspondingly, in the current study, we find a significant increase of auditory and somatosensory MMN with the length of the preceding stimulus train as well as a comparable tendency for the vMMN. Moreover, we show that this increase was driven by a reciprocal negative modulation of deviant and positive modulation of standard responses, suggesting a combined influence of repetition dependent change detection and dynamics akin to stimulus adaptation.

The observed positive modulation of standard responses, particularly in the auditory modality, is in line with the repetition positivity account of Baldeweg and colleagues (Baldeweg, 2006; Baldeweg, 2007; Baldeweg et al., 2004; Haenschel et al., 2005). In the auditory modality, repetition positivity has been isolated as a positive slow wave that accounts for repetition-dependent increases of auditory ERPs up to the P2 component (Haenschel et al., 2005). With regard to its functional role, it has been argued to reflect auditory sensory memory trace formation (Baldeweg et al., 2004; Costa-Faidella, Baldeweg, et al., 2011; Costa-Faidella, Grimm, et al., 2011). Interestingly, MMN studies using the oddball paradigm often report an increasing MMN with standard repetition without further dissecting the contributions from standard and deviant dynamics. A contribution of the standard repetition positivity appears to be particularly dominant in roving stimulus paradigms (Cooper et al., 2013), potentially because a memory trace of the standard stimulus identity must be reestablished after each change of roles for standard and deviant stimuli. It has even been suggested that the memory trace dynamics of the standard observed in response to roving oddball sequences might in fact be the primary driver of train length effects on MMN amplitudes (Baldeweg et al., 2004; Costa-Faidella, Baldeweg, et al., 2011; Costa-Faidella, Grimm, et al., 2011; Haenschel et al., 2005). Importantly, although some evidence exists to suggest an additional role for train length dependent deviant modulation also in roving paradigms (Cowan et al., 1993; Haenschel et al., 2005), a dissection of combined standard and deviant contributions as performed here is rarely described.

Similar to the aMMN, we found the sMMN to be modulated by stimulus repetition. An early repetition positivity effect in the responses to standards was observed prior to 100 ms indicating comparable sensory adaptation dynamics as described for the aMMN. Subsequently, the negative deviant and sMMN responses increase with repetition around the N140 (i.e., around the sMMN peak). While somatosensory deviant responses have previously been shown to decrease with increasing stimulus probability (Akatsuka, Wasaka, Nakata, Kida, & Kakigi, 2007), only few other studies have reported sensitivity of the sMMN to stimulus repetition. Interestingly, in our previous study on somatosensory MMRs (Gijsen et al., 2021) we report the same reciprocal pattern found here: Negative modulation of the deviant and positive modulation of the standard response which result in an increase of the sMMN amplitude with stimulus train length.

In the visual modality, a comparable train length effect to auditory and somatosensory modalities was observed but did not reach statistical significance in the vMMN time window. Given the overall weaker response in the current study for vMMN this might not be surprising. Moreover, discussions about the repetition modulation of vMMN responses are often based on findings concerning the auditory system rather than direct findings in the visual modality. While sensory adaptation to stimulus repetition is generally found throughout the visual system (e.g., Clifford et al., 2007; Grill-Spector et al., 2006) it is rarely directly reported in vMMN studies (but see Kremlacek et al., 2016). Overall, the vMMN literature seems to suggest that the vMMN may

be a rather unstable phenomenon. In fact, by controlling for confounding effects, one study has called the existence of the vMMN for low level features such as the ones used here into question entirely (Male et al., 2020). The vMMN appears to show a much less pronounced spatiotemporal pattern than auditory and somatosensory equivalents, which is reflected in larger variance in the reported topographies and time windows in studies investigating vMMN (but see Section 11 for a discussion of alternative explanations regarding the current study).

4.4 | MMN as a signature of predictive processing

Recent research supports the view that Bayesian perceptual learning mechanisms underlie the generation of mismatch responses such as the MMN (Friston, 2005, 2010; Garrido, Kilner, Stephan, & Friston, 2009). Given the proposal of Bayesian inference and predictive processing as universal principles of perception and perceptual learning in the brain (Friston, 2005, 2010), comparable mismatch responses are expected to be found across sensory modalities. Evidence for the predictive nature of mismatch responses, akin to key findings from the auditory modality, is for instance given by studies showing somatosensory (Andersen & Lundqvist, 2019; Naeije et al., 2018) and visual (Czigler et al., 2006; Kok et al., 2014) MMN in response to predicted but omitted stimuli. Moreover, Ostwald et al. (2012) and Gijsen et al. (2021) have shown that single trial somatosensory MMN and P3 MMRs can be accounted for in terms of surprise signatures of Bayesian inference models tracking stimulus transitions. Similarly, the vMMN has been described as a signature of predictive processing (Kimura et al., 2011; Stefanics et al., 2014), signaling prediction error instead of basic change detection (Stefanics et al., 2018).

Correspondingly, we found comparable mismatch signatures in auditory, somatosensory, and visual modalities. The train length effects observed in our study across modalities have previously been related to predictive processing. Repetition positivity in the auditory modality has been interpreted as a reflection of repetition suppression, resulting from fulfilled prediction (Auksztulewicz & Friston, 2016; Baldeweg, 2007; Costa-Faidella, Baldeweg, et al., 2011; Costa-Faidella, Grimm, et al., 2011). A corresponding negative modulation of deviant responses on the other hand, would signal a failure to suppress prediction error after violation of the regularity established by the current stimulus train. Under such a view, longer trains of repetitions lead to higher precision in the probability estimate which in turn results in a scaling of the prediction error in response to prediction violation (Auksztulewicz & Friston, 2016; Friston, 2005; Friston & Kiebel, 2009). In line with these hypotheses, Garrido and colleagues (Garrido et al., 2008; Garrido, Kilner, Kiebel, & Friston, 2009) used dynamic causal modeling (DCM) to show that the MMN elicited in a roving stimulus paradigm is best explained by the combined dynamics of auditory adaptation and model adjustment. Their network, proposed to underlie MMN generation, was set up as an implementation of hierarchical predictive processing involving bottom-up signals from auditory cortex and top-down modulations by

inferior frontal cortex. Similarly, another DCM study proposed a predictive coding model of pain processing in response to somatosensory oddball sequences, highlighting the role of inferior frontal cortex in top-down modulations of somatosensory potentials (Fardo et al., 2017). As we find involvement of such modality specific sensory and modality independent frontal areas for MMN responses across modalities, our results suggest comparable roles for these sources in a predictive hierarchy.

4.5 | P3 Mismatch responses reflect cross-modal processing

In addition to the modality specific MMN responses, deviants in all three modalities elicited a late positive mismatch component in the P3 time window. Despite differences in the exact latency and extent of this response between modalities, we identified a common mismatch cluster from 300 to 350 ms in central electrodes, followed by a slightly more posterior cluster extending from 400 to 600 ms. Particularly the earlier cluster may correspond to the well-known P3a response, which peaks at around 300 ms after change-onset at (fronto-) central electrodes and is thought to be elicited regardless of sensory modality (Escera et al., 2000; Friedman et al., 2001; Knight & Scabini, 1998; Polich, 2007; Schroger, 1996).

The P3a is closely related to the MMN as they are both elicited during active and passive perception of repeated stimuli interrupted by infrequent stimulus deviations (Polich, 2007; Schroger et al., 2015). While the P3a has been initially related to attentional switches to task-irrelevant but salient stimulus features (Escera et al., 2000; Friedman et al., 2001; Polich, 2007), more recent accounts suggest that the MMN and P3a might reflect two stages of a predictive hierarchy, each representing (potentially differentiable) prediction error responses (Schroger et al., 2015; Wacongne et al., 2011). Similar to the MMN, P3 responses are known to be modulated by stimulus probability (Duncan-Johnson & Donchin, 1977) and can be elicited by unexpected stimulus repetitions (Duncan et al., 2009; Squires et al., 1975) and omissions of predicted sound stimuli (Prete et al.; Sutton et al., 1967), which provides compelling evidence for a role of the P3 in predictive processing. Similar to the MMN responses described above, we found the individual P3 MMR responses in all three modalities to show reciprocal modulations of standards and deviants by stimulus repetition, which has previously only been reported for the auditory modality (Bendixen et al., 2007). This sensitivity to stimulus repetition of mismatch responses in early and late time-windows has been interpreted in terms of regularity and rule extraction in the auditory modality (Bendixen et al., 2007) and is in line with an account of repetition suppression over and above early sensory adaptation.

The MMN and P3 MMR have been shown to be differentially modulated by higher order predictability. The P3 is reduced by the presentation of visual cues preceding an auditory deviant, while the MMN is not affected by the same top-down predictability (Horvath & Bendixen, 2012; Ritter et al., 1999; Sussman et al., 2003). Similarly,

explicit top-down knowledge of sequence regularities has been shown to reduce the P3, while leaving the MMN unaffected (Max et al., 2015). It has thus been suggested that the P3 reflects a higher-level deviance detection system concerned with the significance of the stimulus in providing new information for the system (Horvath et al., 2008). Interestingly, a recent study investigating mismatch responses to different auditory features showed that while the MMN response in an earlier (classical) time window was generally affected by regularity violations, only the later response (P3 range) contained information about the specific features that were violated (An et al., 2021). Furthermore, computational studies indicate that P3 responses reflect specific quantities of unexpectedness as well as updates to a prior belief (Jepma et al., 2016; Kolossa et al., 2015).

Overall, current research provides evidence for the view that the MMN reflects prediction errors at earlier hierarchical stages, primarily concerned with more local regularity extraction, whereas P3 responses reflect more global rule violations which require a certain level of abstraction and information integration (Bekinschtein et al., 2009; Wacongne et al., 2011; Winkler et al., 2005). Our findings of a sensitivity of the P3 response to cross-modal predictive information carried by the multimodal configuration of the stimulus sequence further supports such a view. Across modalities, we found an increased P3 response to mispredicted compared to predicted or unpredictable stimuli, regardless of their role as standards or deviants. Generally, the P3 deviant response in the current study likely reflects a (unsigned) prediction error to a local regularity established by stimulus repetition. However, increased P3 responses to mispredicted stimuli indicate additional violations of global, cross-modal predictions which are extracted from multimodal context information.

The observed pattern suggests influences of precision weighting on prediction errors (Friston & Kiebel, 2009). In case of both predicted and mispredicted stimuli, the cross-modal predictive context allows for more precise predictions (i.e., high prior precision) than in case of the unpredictable stimuli (low prior precision). Under such an interpretation, the precision for mispredicted deviants is high, resulting in a pronounced prediction error response. Since the precision for predicted deviants is also high, the resulting prediction error response is low because the stimulus was suppressed. Even though the size of prediction error to unpredictable deviants could generally be expected in between those of predicted and mispredicted deviants, the observed response is low (similar to that of a predicted deviant), because the prior precision in this context is low. This interpretation is in line with the fact that no significant difference was found between predicted and unpredictable deviants. A similar modulation of multimodal predictability is found for the P3 response to standards. However, interestingly, in case of the standards, the response to predicted stimuli is significantly lower than to unpredictable stimuli. This difference between standards and deviants could be due to the fact that deviants are generally surprising, even if they are more predictable in terms of their cross-modal configuration. Standards, on the other hand, are generally predicted to occur (high precision) which might result in a pronounced suppression of prediction error in case they are additionally cross-modally predicted.

The interpretation of the common P3 cluster as a cross-modal P3a response sensitive to multimodal predictive information is further supported by our source localization results, which particularly indicate prefrontal regions such as the medial frontal, inferior frontal and anterior cingulate cortex as sources of the P3 MMR. Although notoriously diverse, previous research on P3 sources has identified a fronto-parietal network of generators, particularly highlighting the role of prefrontal and anterior cingulate regions in generating the P3 novelty response (P3a; Linden, 2005; Polich, 2007), whereas parietal regions are presumed to be more involved in task-related P3b responses. The identified sources have been shown to be involved in a fronto-parietal network relevant for the supramodal processing of stimulus transitions and deviance detection (Downar et al., 2000; Huang et al., 2005). Similarly, a fronto-parietal attention network (Corbetta & Shulman, 2002) has been shown to be involved in oddball processing in the auditory and visual modalities (Kim, 2014). The network consists of two functionally and anatomically distinct parts which closely interact (Vossel et al., 2014). While the dorsal part of the network is believed to be involved in the allocation of top-down, endogenous attention (e.g., triggered by predictive information), the ventral part is involved in bottom-up, exogenous attention allocation and thus, processing of unexpected stimuli. Importantly, it has been shown that this network operates supramodally to facilitate processing of information from multimodal events (Macaluso, 2010; Macaluso & Driver, 2005). Thus, the predictive information in the multimodal sequences presented in the current study may be processed in such a fronto-parietal network to aid the perception of multimodal stimulus streams. Future research would benefit from studies further investigating such multimodal probabilistic sequences with higher spatial resolution to inform these proposed interpretations.

4.6 | Modeling single-trial EEG responses as signatures of bayesian inference

Given the results of the average- and GLM-based EEG analyses, we aimed to test if the observed modulations of standards, deviants and MMRs by local (train length) and global (cross-modal predictability) sequence properties could be captured by signatures of Bayesian inference. To this end, we compared a simple TLCD model to families of BL models capturing different aspects of the sequence statistics. In light of the literature discussed above, we hypothesized that BL models would outperform the TLCD model in explaining the recorded mismatch responses.

Overall, the BL models outperformed the static TLCD model in all electrodes in the MMN and P3 clusters indicating that these responses reflect dynamics beyond the basic repetition effects observed in the ERP analyses. This result provides evidence to suggest that the MMN and P3 MMR capture the trial-to-trial dynamics of Bayesian inference and are thus markers of probabilistic sequence processing in the brain.

Within the family of BL models, we found that a cross-modally informed model (UCM), tracking cross-modal conditional

dependencies between modalities in addition to unimodal transitions, outperformed a purely unimodal transition probability model (UM) at central electrodes within an early and a late time-window. The cross-modal effects in the late time-window are directly in line with the sensitivity of the P3 cluster to cross-modal predictability discussed above and support an interpretation of P3 mismatch responses to reflect signatures of cross-modal Bayesian inference. Given that cross-modal learning was not explicitly instructed or task-relevant, the results are compatible with the view that the brain is sensitive to cross-modal information by default (Driver & Noesselt, 2008; Ghazanfar & Schroeder, 2006) and that processing multimodal information might be appropriately captured by Bayesian inference (Kording et al., 2007; Shams & Beierholm, 2022). Interestingly, however, an earlier cross-modal effect was found prior to 300 ms which was not reflected in the GLM results, suggesting that potential modulations of MMN signatures by predictability manifest in the dynamics of single trial surprise signals but not in significant mean differences between predictability conditions. Since the earlier cross-modal effect observed in the modeling results was primarily confined to central and fronto-central electrodes it may be related to activity of the frontal generators of the MMN. As discussed above, the frontal cortex is assumed to be involved in MMN generation (Deouell, 2007) in interaction with hierarchically lower sensory sources and has been hypothesized to form top-down predictions about incoming sensory stimuli (Garrido et al., 2008; Garrido, Kilner, Kiebel, & Friston, 2009; Garrido, Kilner, Stephan, & Friston, 2009). This assumption is further supported by our source reconstruction results which show modality independent frontal generators in addition to sensory specific regions to underlie the MMN in auditory, somatosensory, and visual modalities.

Regarding the surprise read-out functions of the BL models, we find a slight dominance of CS in earlier mismatch signatures prior to 200 ms, while the late clusters tend to reflect BS. This is well in line with our previous study performed in the somatosensory modality (Gijsen et al., 2021) and other studies have similarly reported a reflection of BS in P3 mismatch responses (Kolossa et al., 2015; Mars et al., 2008; Ostwald et al., 2012; Seer et al., 2016). Given their differences in reading out the probability estimates of the Bayesian observer, the different surprise signatures in the MMN and P3 MMR might provide some insight into their respective computational roles. CS has been categorized as an instantiation of puzzlement surprise (Faraji et al., 2018) reflecting a mismatch between sensory input and internal model belief which is additionally scaled by belief commitment. Low-probability events are thus more surprising if commitment to the belief (of this estimate) is high. BS reflects incorporation of new information, quantifying an update to the generative model and has been categorized as enlightenment surprise (Faraji et al., 2018). Accordingly, the MMN may be considered a marker of prediction error scaled by belief commitment, whereas the P3 may reflect the subsequent update of the predictive model.

Given that the P3 shows sensitivity to cross-modal prediction violation (GLM results) and tends to carry signatures of multimodal inference and model updating (Bayesian modeling results), we suggest that the P3 likely reflects model updates with respect to the multimodal

context. Although unnoticed by participants, given that the statistical regularities changed across experimental runs, the generative model continuously required updating. As such, our results might be reflective of a volatile sensory environment and relate to previous findings which indicate that later MMRs, such as the P3, reflect belief updates about the volatility of the underlying (hidden) statistics governing sensory observations (e.g., Weber et al., 2020; Weber et al., 2022). We leave it for future research to design experiments which are better suited to evaluate these speculations more specifically and more thoroughly.

4.7 | Limitations

Although we gained valuable insights into the commonalities and differences between mismatch responses in different modalities, our study faces certain limitations in its implementation and scope. First, although reports of weak vMMN responses can be found in the literature, an alternative explanation may lie in the stimulation protocol used in the current study. Our visual stimuli consisted of bilateral flash stimuli with two different intensities, which were presented in the periphery of the visual field. Since, to our knowledge, no other study has used visual flash stimuli to elicit vMMN, our results are not directly comparable to previous research. Moreover, due to the retinotopic organization of the visual cortex (Horton & Hoyt, 1991; Sereno et al., 1995), a "far peripheral" placement (i.e., >60°; Strasburger et al., 2011) of the LED's results in the activation of (primary) visual areas folded deep inside the cortex, in the calcarine sulcus between the hemispheres. It is therefore possible that the visual mismatch responses were not weaker per se but were merely harder to detect by means of EEG.

Further, our results concerning the comparison of the surprise read-out functions provide some indications of the computational roles for early and late MMRs, which are in line with previous research. However, it should be noted that the current study was not specifically designed to investigate their (nuanced) differences. The inclusion of three read-out functions primarily served the purpose of avoiding bias in the comparison of the BL models by prior choice of the read-out. To this end, the most prominent surprise read-out functions used in the literature were included. Research would benefit from future studies specifically designed to compare different surprise measures without the manipulation of other aspects of the underlying models. A valuable overview and suggested experiments for that purpose have been recently provided by Modirshanechi et al. (2022).

5 | CONCLUSION

With the current study, we provide evidence for modality specific and modality independent aspects of mismatch responses in audition, somatosensation, and vision resulting from a simultaneous stream of tri-modal roving stimulus sequences. Our results suggest that responses to stimulus transitions in all three modalities are based on an interaction of hierarchically lower, modality specific areas with hierarchically higher, modality independent frontal areas. We show that similar dynamics underlie these mismatch responses which likely reflect predictive processing and Bayesian inference on unimodal and multimodal sensory input streams.

CONFLICT OF INTEREST

The authors declare that no competing interests exist.

DATA AVAILABILITY STATEMENT

The raw EEG data and the code used for analysis will be made freely available in an online repository upon publication of the article.

ORCID

Miro Grundei https://orcid.org/0000-0002-6396-9674

REFERENCES

Akatsuka, K., Wasaka, T., Nakata, H., Inui, K., Hoshiyama, M., & Kakigi, R. (2005). Mismatch responses related to temporal discrimination of somatosensory stimulation. *Clinical Neurophysiology*, *116*(8), 1930–1937. https://doi.org/10.1016/j.clinph.2005.04.021

Akatsuka, K., Wasaka, T., Nakata, H., Kida, T., Hoshiyama, M., Tamura, Y., & Kakigi, R. (2007). Objective examination for two-point stimulation using a somatosensory oddball paradigm: An MEG study. *Clinical Neurophysiology*, *118*(2), 403–411. https://doi.org/10.1016/j.clinph.2006.09.030

Akatsuka, K., Wasaka, T., Nakata, H., Kida, T., & Kakigi, R. (2007). The effect of stimulus probability on the somatosensory mismatch field. *Experimental Brain Research*, *181*(4), 607–614. https://doi.org/10.1007/s00221-007-0958-4

Alain, C., Woods, D. L., & Ogawa, K. H. (1994). Brain indices of automatic pattern processing. *Neuroreport*, *6*(1), 140–144. https://doi.org/10.1097/00001756-199412300-00036

Allen, M., Fardo, F., Dietz, M. J., Hillebrandt, H., Friston, K. J., Rees, G., & Roepstorff, A. (2016). Anterior insula coordinates hierarchical processing of tactile mismatch responses. *NeuroImage*, *127*, 34–43. https://doi.org/10.1016/j.neuroimage.2015.11.030

An, H., Ho Kei, S., Auksztulewicz, R., & Schnupp, J. W. H. (2021). Do auditory mismatch responses differ between acoustic features? *Frontiers in Human Neuroscience*, *15*, 613903. https://doi.org/10.3389/fnhum.2021.613903

Andersen, L. M., & Lundqvist, D. (2019). Somatosensory responses to nothing: An MEG study of expectations during omission of tactile stimulations. *NeuroImage*, *184*, 78–89. https://doi.org/10.1016/j.neuroimage.2018.09.014

Andric, M., Davis, B., & Hasson, U. (2017). Visual cortex signals a mismatch between regularity of auditory and visual streams. NeuroImage, 157, 648–659. https://doi.org/10.1016/j.neuroimage.2017.05.028

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. Trends in Cognitive Sciences, 16(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. Cortex, 80, 125–140. https://doi.org/10.1016/j.cortex.2015.11.024

Baldeweg, T. (2006). Repetition effects to sounds: Evidence for predictive coding in the auditory system. Trends in Cognitive Sciences, 10(3), 93–94. https://doi.org/10.1016/j.tics.2006.01.010

Baldeweg, T. (2007). ERP repetition effects and mismatch negativity generation: A predictive coding perspective. Journal of Psychophysiology, 21(3–4), 204–213.

Baldeweg, T., Klugman, A., Gruzelier, J., & Hirsch, S. R. (2004). Mismatch negativity potentials and cognitive impairment in schizophrenia. Schizophrenia Research, 69(2–3), 203–217. https://doi.org/10.1016/j.schres.2003.09.009

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. Neural Networks, 23(5), 649–666. https://doi.org/10.1016/j.neunet.2009.12.007

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. Neuron, 76(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. Proceedings of the National Academy of Sciences of the United States of America, 106(5), 1672–1677. https://doi.org/10.1073/pnas.0809667106

Bendixen, A., Roeber, U., & Schroger, E. (2007). Regularity extraction and application in dynamic auditory stimulus sequences. Journal of Cognitive Neuroscience, 19(10), 1664–1677. https://doi.org/10.1162/jocn.2007.19.10.1664

Berg, P., & Scherg, M. (1994). A multiple source approach to the correction of eye artifacts. Electroencephalography and Clinical Neurophysiology, 90(3), 229–241. https://doi.org/10.1016/0013-4694(94)90094-9

Besle, J., Fort, A., & Giard, M. H. (2005). Is the auditory sensory memory sensitive to visual information? Experimental Brain Research, 166(3–4), 337–344. https://doi.org/10.1007/s00221-005-2375-x

Bregman, A. S. (1994). Auditory scene analysis: The perceptual organization of sound. MIT Press.

Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. Journal of Vision, 6(5), 554–564. https://doi.org/10.1167/6.5.2

Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. Brain Research Bulletin, 75(6), 753–760. https://doi.org/10.1016/j.brainresbull.2008.01.009

Butler, J. S., Foxe, J. J., Fiebelkorn, I. C., Mercier, M. R., & Molholm, S. (2012). Multisensory representation of frequency across audition and touch: High density electrical mapping reveals early sensory-perceptual coupling. The Journal of Neuroscience, 32(44), 15338–15344. https://doi.org/10.1523/JNEUROSCI.1796-12.2012

Cao, Y., Summerfield, C., Park, H., Giordano, B. L., & Kayser, C. (2019). Causal inference in the multisensory brain. Neuron, 102(5), 1076–1087 e1078. https://doi.org/10.1016/j.neuron.2019.03.043

Cleary, K. M., Donkers, F. C., Evans, A. M., & Belger, A. (2013). Investigating developmental changes in sensory processing: Visual mismatch response in healthy children. Frontiers in Human Neuroscience, 7, 922. https://doi.org/10.3389/fnhum.2013.00922

Clifford, C. W., Webster, M. A., Stanley, G. B., Stocker, A. A., Kohn, A., Sharpee, T. O., & Schwartz, O. (2007). Visual adaptation: Neural, psychological and computational aspects. Vision Research, 47(25), 3125–3131. https://doi.org/10.1016/j.visres.2007.08.023

Cooper, R. J., Atkinson, R. J., Clark, R. A., & Michie, P. T. (2013). Event-related potentials reveal modelling of auditory repetition in the brain. International Journal of Psychophysiology, 88(1), 74–81. https://doi.org/10.1016/j.ijpsycho.2013.02.003

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews. Neuroscience, 3(3), 201–215. https://doi.org/10.1038/nrn755

Costa-Faidella, J., Baldeweg, T., Grimm, S., & Escera, C. (2011). Interactions between "what" and "when" in the auditory system: Temporal predictability enhances repetition suppression. The Journal of Neuroscience, 31(50), 18590–18597. https://doi.org/10.1523/JNEUROSCI.2599-11.2011

Costa-Faidella, J., Grimm, S., Slabu, L., Diaz-Santaella, F., & Escera, C. (2011). Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. Psychophysiology, 48(6), 774–783. https://doi.org/10.1111/j.1469-8986.2010.01144.x

Cowan, N., Winkler, I., Teder, W., & Naatanen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). Journal of Experimental Psychology. Learning, Memory, and Cognition, 19(4), 909–921. https://doi.org/10.1037/0278-7393.19.4.909

Czigler, I., Winkler, I., Pato, L., Varnagy, A., Weisz, J., & Balazs, L. (2006). Visual temporal window of integration as revealed by the visual mismatch negativity event-related potential to stimulus omissions. Brain Research, 1104(1), 129–140. https://doi.org/10.1016/j.brainres.2006.05.034

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The neural representation of sequences: From transition probabilities to algebraic patterns and linguistic trees. Neuron, 88(1), 2–19. https://doi.org/10.1016/j.neuron.2015.09.019

Deouell, L. Y. (2007). The frontal generator of the mismatch negativity revisited. Journal of Psychophysiology, 21(3–4), 188–203.

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. Nature Neuroscience, 3(3), 277–283. https://doi.org/10.1038/72991

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron, 57(1), 11–23. https://doi.org/10.1016/j.neuron.2007.12.013

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Naatanen, R., Polich, J., Reinvang, I., & Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. Clinical Neurophysiology, 120(11), 1883–1908. https://doi.org/10.1016/j.clinph.2009.07.045

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. Psychophysiology, 14(5), 456–467. https://doi.org/10.1111/j.1469-8986.1977.tb01312.x

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage, 25(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Escera, C., Alho, K., Schroger, E., & Winkler, I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiology & Neuro-Otology, 5(3–4), 151–166. https://doi.org/10.1159/000013877

Faraji, M., Preuschoff, K., & Gerstner, W. (2018). Balancing new against old information: The role of puzzlement surprise in learning. Neural Computation, 30(1), 34–83. https://doi.org/10.1162/neco_a_01025

Fardo, F., Auksztulewicz, R., Allen, M., Dietz, M. J., Roepstorff, A., & Friston, K. J. (2017). Expectation violation and attention to pain jointly

modulate neural gain in somatosensory cortex. *NeuroImage*, *153*, 109–121. https://doi.org/10.1016/j.neuroimage.2017.03.041

File, D., File, B., Bodnar, F., Sulykos, I., Kecskes-Kovacs, K., & Cziglers, I. (2017). Visual mismatch negativity (vMMN) for low- and high-level deviances: A control study. *Attention, Perception, & Psychophysics*, *79*(7), 2153–2170. https://doi.org/10.3758/s13414-017-1373-y

Fitzgerald, K., & Todd, J. (2020). Making sense of mismatch negativity. *Frontiers in Psychiatry*, *11*, 468. https://doi.org/10.3389/fpsyt.2020.00468

Flandin, G., & Penny, W. D. (2007). Bayesian fMRI data analysis with sparse spatial basis function priors. *NeuroImage*, *34*(3), 1108–1125. https://doi.org/10.1016/j.neuroimage.2006.10.005

Friedel, E. B. N., Bach, M., & Heinrich, S. P. (2020). Attentional interactions between vision and hearing in event-related responses to Crossmodal and conjunct oddballs. *Multisensory Research*, *33*(3), 251–275. https://doi.org/10.1163/22134808-20191329

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews*, *25*(4), 355–373. https://doi.org/10.1016/s0149-7634(01)00019-7

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Kilner, J. M. (2008). The functional anatomy of the MMN: A DCM study of the roving paradigm. *NeuroImage*, *42*(2), 936–944. https://doi.org/10.1016/j.neuroimage.2008.05.018

Garrido, M. I., Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2009). Dynamic causal modeling of the response to frequency deviants. *Journal of Neurophysiology*, *101*(5), 2620–2631. https://doi.org/10.1152/jn.90291.2008

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463. https://doi.org/10.1016/j.clinph.2008.11.029

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

Gijsen, S., Grundei, M., Lange, R. T., Ostwald, D., & Blankenburg, F. (2021). Neural surprise in somatosensory Bayesian learning. *PLoS Computational Biology*, *17*(2), e1008068. https://doi.org/10.1371/journal.pcbi.1008068

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *290*(1038), 181–197. https://doi.org/10.1098/rstb.1980.0090

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *The Journal of Neuroscience*, *25*(45), 10494–10501. https://doi.org/10.1523/JNEUROSCI.1227-05.2005

Hedge, C., Stothart, G., Todd Jones, J., Rojas Frias, P., Magee, K. L., & Brooks, J. C. (2015). A frontal attention mechanism in the visual mismatch negativity. *Behavioural Brain Research*, *293*, 173–181. https://doi.org/10.1016/j.bbr.2015.07.022

Heilbron, M., & Chait, M. (2018). Great expectations: Is there evidence for predictive coding in auditory cortex? *Neuroscience*, *389*, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Heslenfeld, D. J. (2003). Visual mismatch negativity. In *Detection of change: Event-related potential and fMRI findings* (pp. 41–59). Kluwer Academic Publishers.

Horton, J. C., & Hoyt, W. F. (1991). The representation of the visual field in human striate cortex. A revision of the classic Holmes map. *Archives of Ophthalmology*, *109*(6), 816–824. https://doi.org/10.1001/archopht.1991.01080060080030

Horvath, J., & Bendixen, A. (2012). Preventing distraction by probabilistic cueing. *International Journal of Psychophysiology*, *83*(3), 342–347. https://doi.org/10.1016/j.ijpsycho.2011.11.019

Horvath, J., & Winkler, I. (2004). How the human auditory system treats repetition amongst change. *Neuroscience Letters*, *368*(2), 157–161. https://doi.org/10.1016/j.neulet.2004.07.004

Horvath, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology*, *79*(2), 139–147. https://doi.org/10.1016/j.biopsycho.2008.04.001

Hu, L., Zhao, C., Li, H., & Valentini, E. (2013). Mismatch responses evoked by nociceptive stimuli. *Psychophysiology*, *50*(2), 158–173. https://doi.org/10.1111/psyp.12000

Huang, M. X., Lee, R. R., Miller, G. A., Thoma, R. J., Hanlon, F. M., Paulson, K. M., Martin, K., Harrington, D. L., Weisend, M. P., Edgar, J. C., & Canive, J. M. (2005). A parietal-frontal network studied by somatosensory oddball MEG responses, and its cross-modal consistency. *NeuroImage*, *28*(1), 99–114. https://doi.org/10.1016/j.neuroimage.2005.05.036

Huettel, S. A., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: Dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, *5*(5), 485–490. https://doi.org/10.1038/nn841

Hughes, H. C., Darcey, T. M., Barkan, H. I., Williamson, P. D., Roberts, D. W., & Aslin, C. H. (2001). Responses of human auditory association cortex to the omission of an expected acoustic event. *NeuroImage*, *13*(6 Pt 1), 1073–1089. https://doi.org/10.1006/nimg.2001.0766

Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology*, *19*(2), 113–124. https://doi.org/10.1097/00004691-200203000-00002

Imada, T., Hari, R., Loveless, N., McEvoy, L., & Sams, M. (1993). Determinants of the auditory mismatch response. *Electroencephalography and Clinical Neurophysiology*, *87*(3), 144–153. https://doi.org/10.1016/0013-4694(93)90120-k

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007

Jaaskelainen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levanen, S., Lin, F. H., May, P., Melcher, J., Stufflebeam, S., Tiitinen, H., & Belliveau, J. W. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(17), 6809–6814. https://doi.org/10.1073/pnas.0303760101

Javitt, D. C., Grochowski, S., Shelley, A. M., & Ritter, W. (1998). Impaired mismatch negativity (MMN) generation in schizophrenia as a function of stimulus deviance, probability, and interstimulus/interdeviant interval. *Electroencephalography and Clinical Neurophysiology*, *108*(2), 143–153. https://doi.org/10.1016/s0168-5597(97)00073-7

Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic regulation of learning rate

in a dynamic environment. *PLoS Computational Biology*, *12*(10), e1005171. https://doi.org/10.1371/journal.pcbi.1005171

Kekoni, J., Hamalainen, H., Saarinen, M., Grohn, J., Reinikainen, K., Lehtokoski, A., & Naatanen, R. (1997). Rate effect and mismatch responses in the somatosensory system: ERP-recordings in humans. *Biological Psychology*, *46*(2), 125–142. https://doi.org/10.1016/s0301-0511(97)05249-6

Kiat, J. E. (2018). Assessing cross-modal target transition effects with a visual-auditory oddball. *International Journal of Psychophysiology*, *129*, 58–66. https://doi.org/10.1016/j.ijpsycho.2018.04.010

Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2005). Applications of random field theory to electrophysiology. *Neuroscience Letters*, *374*(3), 174–178. https://doi.org/10.1016/j.neulet.2004.10.052

Kim, H. (2014). Involvement of the dorsal and ventral attention networks in oddball stimulus processing: A meta-analysis. *Human Brain Mapping*, *35*(5), 2265–2284. https://doi.org/10.1002/hbm.22326

Kimura, M., Kondo, H., Ohira, H., & Schroger, E. (2012). Unintentional temporal context-based prediction of emotional faces: An electrophysiological study. *Cerebral Cortex*, *22*(8), 1774–1785. https://doi.org/10.1093/cercor/bhr244

Kimura, M., Ohira, H., & Schroger, E. (2010). Localizing sensory and cognitive systems for pre-attentive visual deviance detection: An sLORETA analysis of the data of Kimura et al. (2009). *Neuroscience Letters*, *485*(3), 198–203. https://doi.org/10.1016/j.neulet.2010.09.011

Kimura, M., Schroger, E., & Czigler, I. (2011). Visual mismatch negativity and its importance in visual cognitive sciences. *Neuroreport*, *22*(14), 669–673. https://doi.org/10.1097/WNR.0b013e32834973ba

Knight, R. T., & Scabini, D. (1998). Anatomic bases of event-related potentials and their relationship to novelty detection in humans. *Journal of Clinical Neurophysiology*, *15*(1), 3–13. https://doi.org/10.1097/00004691-199801000-00003

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience*, *26*(7), 1546–1554. https://doi.org/10.1162/jocn_a_00562

Kolossa, A., Kopp, B., & Fingscheidt, T. (2015). A computational analysis of the neural bases of Bayesian inference. *NeuroImage*, *106*, 222–237. https://doi.org/10.1016/j.neuroimage.2014.11.007

Kording, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS One*, *2*(9), e943. https://doi.org/10.1371/journal.pone.0000943

Kremlacek, J., Kreegipuu, K., Tales, A., Astikainen, P., Poldver, N., Naatanen, R., & Stefanics, G. (2016). Visual mismatch negativity (vMMN): A review and meta-analysis of studies in psychiatric and neurological disorders. *Cortex*, *80*, 76–112. https://doi.org/10.1016/j.cortex.2016.03.017

Kuldkepp, N., Kreegipuu, K., Raidvee, A., Naatanen, R., & Allik, J. (2013). Unattended and attended visual change detection of motion as indexed by event-related potentials and its behavioral correlates. *Frontiers in Human Neuroscience*, *7*, 476. https://doi.org/10.3389/fnhum.2013.00476

Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J., & Stephan, K. E. (2013). Modelling trial-by-trial changes in the mismatch negativity. *PLoS Computational Biology*, *9*(2), e1002911. https://doi.org/10.1371/journal.pcbi.1002911

Linden, D. E. (2005). The p300: Where in the brain is it produced and what does it tell us? *The Neuroscientist*, *11*(6), 563–576. https://doi.org/10.1177/1073858405280524

Macaluso, E. (2010). Orienting of spatial attention and the interplay between the senses. *Cortex*, *46*(3), 282–297. https://doi.org/10.1016/j.cortex.2009.05.010

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences*, *28*(5), 264–271. https://doi.org/10.1016/j.tins.2005.03.008

Macdonald, M., & Campbell, K. (2011). Effects of a violation of an expected increase or decrease in intensity on detection of change within an auditory pattern. *Brain and Cognition*, *77*(3), 438–445. https://doi.org/10.1016/j.bandc.2011.08.014

Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multi-scale process of sequence learning in humans. *eLife*, *8*, e41541. https://doi.org/10.7554/eLife.41541

Male, A. G., O'Shea, R. P., Schroger, E., Muller, D., Roeber, U., & Widmann, A. (2020). The quest for the genuine visual mismatch negativity (vMMN): Event-related potential indications of deviance detection for low-level visual features. *Psychophysiology*, *57*(6), e13576. https://doi.org/10.1111/psyp.13576

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

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *The Journal of Neuroscience*, *28*(47), 12539–12545. https://doi.org/10.1523/JNEUROSCI.2925-08.2008

Max, C., Widmann, A., Schroger, E., & Sussman, E. (2015). Effects of explicit knowledge and predictability on auditory distraction and target performance. *International Journal of Psychophysiology*, *98*(2 Pt 1), 174–181. https://doi.org/10.1016/j.ijpsycho.2015.09.006

May, P., Tiitinen, H., Ilmoniemi, R. J., Nyman, G., Taylor, J. G., & Naatanen, R. (1999). Frequency change detection in human auditory cortex. *Journal of Computational Neuroscience*, *6*(2), 99–120. https://doi.org/10.1023/a:1008896417606

May, P. J., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*(1), 66–122. https://doi.org/10.1111/j.1469-8986.2009.00856.x

Modirshanechi, A., Brea, J., & Gerstner, W. (2022). A taxonomy of surprise definitions. *Journal of Mathematical Psychology*, *110*, 102712.

Molholm, S., Martinez, A., Ritter, W., Javitt, D. C., & Foxe, J. J. (2005). The neural circuitry of pre-attentive auditory change-detection: An fMRI study of pitch and duration mismatch negativity generators. *Cerebral Cortex*, *15*(5), 545–551. https://doi.org/10.1093/cercor/bhh155

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*(2), 201–233.

Naatanen, R., Gaillard, A. W., & Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313–329. https://doi.org/10.1016/0001-6918(78)90006-9

Naatanen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, *42*(1), 25–32. https://doi.org/10.1111/j.1469-8986.2005.00256.x

Naatanen, R., & Näätänen, R. (1992). *Attention and brain function*. Psychology Press.

Naatanen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026

Naeije, G., Vaulet, T., Wens, V., Marty, B., Goldman, S., & De Tiege, X. (2016). Multilevel cortical processing of somatosensory novelty: A magnetoencephalography study. *Frontiers in Human Neuroscience*, *10*, 259. https://doi.org/10.3389/fnhum.2016.00259

Naeije, G., Vaulet, T., Wens, V., Marty, B., Goldman, S., & De Tiege, X. (2018). Neural basis of early somatosensory change detection: A magnetoencephalography study. *Brain Topography*, *31*(2), 242–256. https://doi.org/10.1007/s10548-017-0591-x

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y., & Schroger, E. (2002). Differential contribution of frontal and temporal cortices to

auditory change detection: fMRI and ERP results. *NeuroImage*, *15*(1), 167–174. https://doi.org/10.1006/nimg.2001.0970

Ossmy, O., Moran, R., Pfeffer, T., Tsetsos, K., Usher, M., & Donner, T. H. (2013). The timescale of perceptual evidence integration can be adapted to the environment. *Current Biology*, *23*(11), 981–986. https://doi.org/10.1016/j.cub.2013.04.039

Ostwald, D., Spitzer, B., Guggenmos, M., Schmidt, T. T., Kiebel, S. J., & Blankenburg, F. (2012). Evidence for neural encoding of Bayesian surprise in human somatosensation. *NeuroImage*, *62*(1), 177–188. https://doi.org/10.1016/j.neuroimage.2012.04.050

Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *International Journal of Psychophysiology*, *88*(2), 109–123. https://doi.org/10.1016/j.ijpsycho.2013.03.015

Parmentier, F. B., Elsley, J. V., Andres, P., & Barcelo, F. (2011). Why are auditory novels distracting? Contrasting the roles of novelty, violation of expectation and stimulus change. *Cognition*, *119*(3), 374–380. https://doi.org/10.1016/j.cognition.2011.02.001

Pazo-Alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: A review. *Biological Psychology*, *63*(3), 199–236. https://doi.org/10.1016/s0301-0511(03)00049-8

Penny, W., Kiebel, S., & Friston, K. (2003). Variational Bayesian inference for fMRI time series. *NeuroImage*, *19*(3), 727–741. https://doi.org/10.1016/s1053-8119(03)00071-5

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., & Leff, A. P. (2010). Comparing families of dynamic causal models. *PLoS Computational Biology*, *6*(3), e1000709. https://doi.org/10.1371/journal.pcbi.1000709

Penny, W. D., Trujillo-Barreto, N. J., & Friston, K. J. (2005). Bayesian fMRI time series analysis with spatial priors. *NeuroImage*, *24*(2), 350–362. https://doi.org/10.1016/j.neuroimage.2004.08.034

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Prete, D. A., Heikoop, D., McGillivray, J. E., Reilly, J. P., & Trainor, L. J. (2022). The sound of silence: Predictive error responses to unexpected sound omission in adults. *The European Journal of Neuroscience*, *55*, 1972–1985. https://doi.org/10.1111/ejn.15660

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ritter, W., Sussman, E., Deacon, D., Cowan, N., & Vaughan, H. G., Jr. (1999). Two cognitive systems simultaneously prepared for opposite events. *Psychophysiology*, *36*(6), 835–838.

Rohe, T., Ehlis, A. C., & Noppeney, U. (2019). The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nature Communications*, *10*(1), 1907. https://doi.org/10.1038/s41467-019-09664-2

Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biology*, *13*(2), e1002073. https://doi.org/10.1371/journal.pbio.1002073

Runyan, C. A., Piasini, E., Panzeri, S., & Harvey, C. D. (2017). Distinct timescales of population coding across cortex. *Nature*, *548*(7665), 92–96. https://doi.org/10.1038/nature23020

Sabeti, M., Katebi, S. D., Rastgar, K., & Azimifar, Z. (2016). A multi-resolution approach to localize neural sources of P300 event-related brain potential. *Computer Methods and Programs in Biomedicine*, *133*, 155–168. https://doi.org/10.1016/j.cmpb.2016.05.013

Salisbury, D. F. (2012). Finding the missing stimulus mismatch negativity (MMN): Emitted MMN to violations of an auditory gestalt. *Psychophysiology*, *49*(4), 544–548. https://doi.org/10.1111/j.1469-8986.2011.01336.x

Sams, M., Alho, K., & Naatanen, R. (1983). Sequential effects on the ERP in discriminating two stimuli. *Biological Psychology*, *17*(1), 41–58. https://doi.org/10.1016/0301-0511(83)90065-0

Schroger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*, *8*(6), 527–539. https://doi.org/10.1162/jocn.1996.8.6.527

Schroger, E., Marzecova, A., & SanMiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. *The European Journal of Neuroscience*, *41*(5), 641–664. https://doi.org/10.1111/ejn.12816

Seer, C., Lange, F., Boos, M., Dengler, R., & Kopp, B. (2016). Prior probabilities modulate cortical surprise responses: A study of event-related potentials. *Brain and Cognition*, *106*, 78–89. https://doi.org/10.1016/j.bandc.2016.04.011

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*(5212), 889–893. https://doi.org/10.1126/science.7754376

Shams, L., & Beierholm, U. (2022). Bayesian causal inference: A unifying neuroscience theory. *Neuroscience and Biobehavioral Reviews*, *137*, 104619. https://doi.org/10.1016/j.neubiorev.2022.104619

Shams, L., Wozny, D. R., Kim, R., & Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in Psychology*, *2*, 264. https://doi.org/10.3389/fpsyg.2011.00264

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Shen, G., Smyk, N. J., Meltzoff, A. N., & Marshall, P. J. (2018). Using somatosensory mismatch responses as a window into somatotopic processing of tactile stimulation. *Psychophysiology*, *55*(5), e13030. https://doi.org/10.1111/psyp.13030

Shinozaki, N., Yabe, H., Sutoh, T., Hiruma, T., & Kaneko, S. (1998). Somatosensory automatic responses to deviant stimuli. *Brain Research. Cognitive Brain Research*, *7*(2), 165–171. https://doi.org/10.1016/s0926-6410(98)00020-2

Spackman, L. A., Boyd, S. G., & Towell, A. (2007). Effects of stimulus frequency and duration on somatosensory discrimination responses. *Experimental Brain Research*, *177*(1), 21–30. https://doi.org/10.1007/s00221-006-0650-0

Spackman, L. A., Towell, A., & Boyd, S. G. (2010). Somatosensory discrimination: An intracranial event-related potential study of children with refractory epilepsy. *Brain Research*, *1310*, 68–76. https://doi.org/10.1016/j.brainres.2009.10.072

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401. https://doi.org/10.1016/0013-4694(75)90263-1

Stefanics, G., Heinzle, J., Horvath, A. A., & Stephan, K. E. (2018). Visual mismatch and predictive coding: A computational single-trial ERP study. *The Journal of Neuroscience*, *38*(16), 4020–4030. https://doi.org/10.1523/JNEUROSCI.3365-17.2018

Stefanics, G., Kremlacek, J., & Czigler, I. (2014). Visual mismatch negativity: A predictive coding view. *Frontiers in Human Neuroscience*, *8*, 666. https://doi.org/10.3389/fnhum.2014.00666

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004–1017. https://doi.org/10.1016/j.neuroimage.2009.03.025

Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, *11*(5), 13. https://doi.org/10.1167/11.5.13

Sussman, E., Ritter, W., & Vaughan, H. G., Jr. (1998). Predictability of stimulus deviance and the mismatch negativity. *Neuroreport*, *9*(18), 4167–4170. https://doi.org/10.1097/00001756-199812210-00031

Sussman, E., Winkler, I., & Schroger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic*

Bulletin & Review, *10*(3), 630–637. https://doi.org/10.3758/bf03196525

Sussman, E. S. (2005). Integration and segregation in auditory scene analysis. *The Journal of the Acoustical Society of America*, *117*(3 Pt 1), 1285–1298. https://doi.org/10.1121/1.1854312

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188. https://doi.org/10.1126/science.150.3700.1187

Sutton, S., Tueting, P., Zubin, J., & John, E. R. (1967). Information delivery and the sensory evoked potential. *Science*, *155*(3768), 1436–1439. https://doi.org/10.1126/science.155.3768.1436

Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *The Journal of Neuroscience*, *24*(46), 10440–10453. https://doi.org/10.1523/JNEUROSCI.1905-04.2004

Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, *6*(4), 391–398. https://doi.org/10.1038/nn1032

Urakawa, T., Inui, K., Yamashiro, K., & Kakigi, R. (2010). Cortical dynamics of the visual change detection process. *Psychophysiology*, *47*(5), 905–912. https://doi.org/10.1111/j.1469-8986.2010.00987.x

Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik* (Vol. 9). Voss.

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *The Neuroscientist*, *20*(2), 150–159. https://doi.org/10.1177/1073858413494269

Wacongne, C., Changeux, J. P., & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *The Journal of Neuroscience*, *32*(11), 3665–3678. https://doi.org/10.1523/JNEUROSCI.5003-11.2012

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(51), 20754–20759. https://doi.org/10.1073/pnas.1117807108

Weber, L. A., Diaconescu, A. O., Mathys, C., Schmidt, A., Kometer, M., Vollenweider, F., & Stephan, K. E. (2020). Ketamine affects prediction errors about statistical regularities: A computational single-trial analysis of the mismatch negativity. *The Journal of Neuroscience*, *40*(29), 5658–5668. https://doi.org/10.1523/JNEUROSCI.3069-19.2020

Weber, L. A., Tomiello, S., Schobi, D., Wellstein, K. V., Mueller, D., Iglesias, S., & Stephan, K. E. (2022). Auditory mismatch responses are differentially sensitive to changes in muscarinic acetylcholine versus dopamine receptor function. *eLife*, *11*, e74835. https://doi.org/10.7554/eLife.74835

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(3–4), 147–163.

Winkler, I., & Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, *83*(2), 132–143. https://doi.org/10.1016/j.ijpsycho.2011.10.001

Winkler, I., Czigler, I., Sussman, E., Horvath, J., & Balazs, L. (2005). Preattentive binding of auditory and visual stimulus features. *Journal of Cognitive Neuroscience*, *17*(2), 320–339. https://doi.org/10.1162/0898929053124866

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*(12), 532–540. https://doi.org/10.1016/j.tics.2009.09.003

Woolrich, M. W. (2012). Bayesian inference in FMRI. *NeuroImage*, *62*(2), 801–810. https://doi.org/10.1016/j.neuroimage.2011.10.047

Yabe, H., Tervaniemi, M., Reinikainen, K., & Naatanen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *Neuroreport*, *8*(8), 1971–1974. https://doi.org/10.1097/00001756-199705260-00035

Yucel, G., McCarthy, G., & Belger, A. (2007). fMRI reveals that involuntary visual deviance processing is resource limited. *NeuroImage*, *34*(3), 1245–1252. https://doi.org/10.1016/j.neuroimage.2006.08.050

Zhao, C., Valentini, E., & Hu, L. (2015). Functional features of crossmodal mismatch responses. *Experimental Brain Research*, *233*(2), 617–629. https://doi.org/10.1007/s00221-014-4141-4