biology

MDPI

Article

Statistical Learning of Incidental Perceptual Regularities Induces Sensory Conditioned Cortical Responses

Antonino Greco [1,2,3,*], Marco D'Alessandro [4], Giuseppe Gallitto [5], Clara Rastelli [3,6], Christoph Braun [1,2,3] and Andrea Caria [6,*]

1 Department of Neural Dynamics and Magnetoencephalography, Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany
2 Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, 72076 Tübingen, Germany
3 MEG Center, University of Tübingen, 72076 Tübingen, Germany
4 Institute of Cognitive Sciences and Technologies, National Research Council, 00185 Rome, Italy
5 Department of Neurology, University Hospital Essen, 45147 Essen, Germany
6 Department of Psychology and Cognitive Science, University of Trento, 38068 Rovereto, Italy
* Correspondence: antonino.greco@uni-tuebingen.de (A.G.); andrea.caria@unitn.it (A.C.)

Simple Summary: Our study demonstrated neural encoding of incidental sensory regularities leading to modulation of cortical responses to both predictive and predicted sensory stimuli. As in the case of goal-directed behavior, such task-irrelevant predictive mechanisms might result from the brain's intrinsic drive of reducing uncertainty about state transition dynamics of the environment.

Abstract: Statistical learning of sensory patterns can lead to predictive neural processes enhancing stimulus perception and enabling fast deviancy detection. Predictive processes have been extensively demonstrated when environmental statistical regularities are relevant to task execution. Preliminary evidence indicates that statistical learning can even occur independently of task relevance and top-down attention, although the temporal profile and neural mechanisms underlying sensory predictions and error signals induced by statistical learning of incidental sensory regularities remain unclear. In our study, we adopted an implicit sensory conditioning paradigm that elicited the generation of specific perceptual priors in relation to task-irrelevant audio–visual associations, while recording Electroencephalography (EEG). Our results showed that learning task-irrelevant associations between audio–visual stimuli resulted in anticipatory neural responses to predictive auditory stimuli conveying anticipatory signals of expected visual stimulus presence or absence. Moreover, we observed specific modulation of cortical responses to probabilistic visual stimulus presentation or omission. Pattern similarity analysis indicated that predictive auditory stimuli tended to resemble the response to expected visual stimulus presence or absence. Remarkably, Hierarchical Gaussian filter modeling estimating dynamic changes of prediction error signals in relation to differential probabilistic occurrences of audio–visual stimuli further demonstrated instantiation of predictive neural signals by showing distinct neural processing of prediction error in relation to violation of expected visual stimulus presence or absence. Overall, our findings indicated that statistical learning of non-salient and task-irrelevant perceptual regularities could induce the generation of neural priors at the time of predictive stimulus presentation, possibly conveying sensory-specific information about the predicted consecutive stimulus.

Keywords: sensory conditioning; associative learning; predictive processing; statistical learning; perceptual regularities; hierarchical gaussian filter; pattern similarity analysis; EEG

check for updates

Citation: Greco, A.; D'Alessandro, M.; Gallitto, G.; Rastelli, C.; Braun, C.; Caria, A. Statistical Learning of Incidental Perceptual Regularities Induces Sensory Conditioned Cortical Responses. *Biology* **2024**, *13*, 576. https://doi.org/10.3390/biology13080576

Received: 29 May 2024
Revised: 24 July 2024
Accepted: 29 July 2024
Published: 30 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Human sensory perception is posited to depend upon the rapid encoding of probabilistic occurrences of perceptual information and the consequent timely generation of specific sensory priors [1–7]. A number of studies indicated that statistical learning of sensory

patterns leads to predictive neural processes, enhancing visual perception and enabling fast deviancy detection [8–15]. Predictive models of perception postulate that high-level generative neural signals provide anticipatory neural representations of sensory signals, reducing perceptual surprisal, and minimizing computational effort [3,16–19]. Accordingly, rapid learning of relevant perceptual regularities has been shown to induce attenuation of neural population reaction to predictable sensory stimuli, as well as enhanced responses to unpredictable information [9,20–25]. The neural expectation suppression effect [22] is proposed to arise either from the dampening of the stimulus response as a result of a decrease in global surprise signals [26,27] or from a sharpening of cortical response to specific sensory input and an inherent reduction of prediction errors [10,28]. In line with this last assumption, prior expectations can facilitate sensory processing by increasing fine-tuning of the primary sensory cortex [10,21,28,29]. Moreover, the encoding of predictable sensory patterns was also associated with anticipatory perceptual processes such as pre-stimulus-specific baseline shifts [30,31], and the pre-activation of the primary sensory cortex immediately before stimulus presentation, resulting in timely cortical instantiation of the expected representational content for the stimulus [32–34], as well as attenuation of alpha-band oscillations and the contingent negative variation [35]. On the other hand, expectation-related effects are still under debate, and in some statistical learning studies, these effects do not emerge when checking for potential confounding factors such as repetition suppression, adaptation, and novelty effects [36–41].

In addition, learned perceptual statistical regularities and resulting predictive neural processes are significantly influenced by task relevance and attention [32,42–46]. However, statistical learning in the absence of explicit top-down attention can also occur, and this leads to an attentional suppression effect [47]. In addition, attenuation of cortical fMRI responses to predicted stimuli was also observed in cases of task-irrelevant sensory input [9,21,48]. Stimulus relevance and probability appear to have dissociable effects on visual processing [14,45]. For instance, stimulus relevance can enhance the precision of stimulus processing by suppressing internal noise, whereas sensory signal probability would bias stimulus detection by increasing the baseline activity of signal-selective units during early visual processing [14]. Notably, the learned probability of sensory stimulus occurrence can significantly and differentially impact both late and early response stages [14,32].

Overall, whether the brain can learn statistical associations of task-irrelevant stimuli remains contentious. Some studies provided evidence supporting this associative learning mechanism [20,21,48], whereas other research suggested that an association of task-irrelevant sensory stimuli does not occur [49–51]. In addition, while the processes associated with responses to predictable information, such as cortical response attenuation, are well documented, the neural mechanisms underlying the postulated anticipatory processing of predictive and predicted task-irrelevant sensory information have yet to be described.

Here, we assumed that the learning of non-salient sensory regularities can be shaped through associative mechanisms that might lead to sensory-conditioned stimuli capable of inducing anticipatory responses similar to unconditioned responses [52–55]. In previous sensory pre-conditioning studies, an initial sensory conditioning phase typically preceded Pavlovian conditioning; this aspect raises the question of whether it was in fact the consecutive conditioning phase involving potent biological stimuli that significantly influenced learned associations between neutral sensory stimuli. On the other hand, functional neuroimaging studies suggested that learned associations of incidental audio–visual regularities can occur independently of stimulus salience and relevance and induce modifications of neural responses to paired sensory inputs [21,29]. Associative effects of such neural responses were explained in terms of predictions and prediction errors of expected sensory regularities [21]., However, as of yet, the temporal and representational aspects of the neural responses underlying predictions and prediction error signals related to learned associations of non-salient and task-irrelevant, yet statistically regular, sensory stimuli have not been elucidated.

We based our hypothesis in our investigation on predictive processing principles. We hypothesized that associative learning of non-relevant, but probabilistically interrelated, neutral sensory stimuli resulted in early neural responses to predictive sensory stimuli associated with anticipatory signals which are informative of the subsequent sensory input, as well as a specific response modulation to the predicted stimulus. To this aim, we collected electroencephalographic (EEG) data during a novel implicit sensory conditioning paradigm that, throughout its unfolding, was expected to induce increasingly specific perceptual priors relative to probabilistic but task-irrelevant audio–visual associations. The experimental protocol entailed participants being exposed to incidental probabilistic associations of non-salient audio–visual patterns while engaged in a main stimulus detection task (Figure 1). We specifically assumed that the implicit learning of task-irrelevant audio–visual associations would result in anticipatory brain activity associated with auditory stimuli (audio period). This anticipatory activity was expected to convey information about the consecutive visual stimulus presentation (post-audio period), possibly being evidenced by the increased similarity between the pattern of neural response to the predictive auditory stimulus and that to the predicted visual stimulus. We also conjectured that such an anticipatory forward mechanism should be differentially implemented in cases of probabilistic visual stimulus presence and absence and would also result in distinct prediction error assessment. In our experimental protocol-learned perceptual associations were assumed to occur exclusively due to repeated exposure to statistical regularity of sensory stimuli. To test our hypotheses, we examined neural activity during sensory conditioning by combining ERP analysis and multivariate classification. We then performed Pattern Similarity Analysis [56,57] to assess the relationship of activation pattern evoked by a predictive auditory stimulus with that evoked by expected visual stimulus presentation or omission, by computing the cross-validated Mahalanobis distance (cvMD) [58]. Finally, we modeled the evolution of prediction error signals during implicit learning of incidental perceptual regularities [11,59,60] using the Hierarchical Gaussian filter (HGF) modeling, a computational approach that was shown to successfully explain associative learning in terms of probabilistic perceptual priors [13]. The HGF, a Bayesian ideal observer model for characterizing inferences of uncertain perceptual inputs [61], allowed us to estimate dynamic changes of prediction error signals in relation to differential probabilistic occurrences of visual input presentation and omission during sensory conditioning. Overall, our study indicates that the human brain rapidly encodes task-irrelevant pairings of sensory stimuli, and it shows that neural responses to learned associations of paired audio–visual stimuli can be explained within the predictive coding framework.

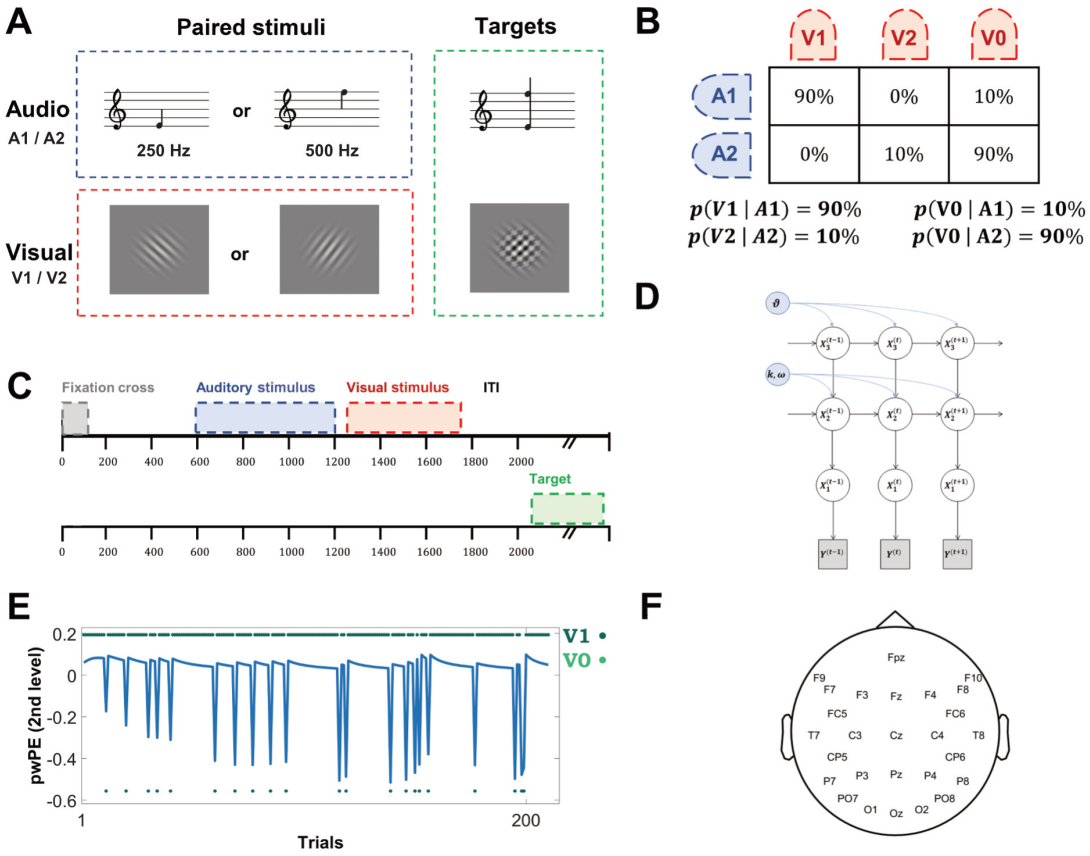


Figure 1. (**A**). The experimental procedure entailed passive exposition to a stream of task-irrelevant auditory stimuli, A1 and A2. These consist of low and high-frequency pure tones, 250 Hz and 500 Hz, respectively (counterbalanced across participants). These were paired with task-irrelevant visual stimuli, V1 and V2, consisting of two white Gabor patches with a 45° and a 135° orientation, presented against a grey background. The main task entailed detecting and pressing a button in response to two specific targets only, consisting of an auditory stimulus combining both A1 and A2 and a visual stimulus combining both V1 and V2. (**B**). Contingency table showing the probabilistic occurrences of visual stimuli presentation given each auditory stimulus, and the resulting conditional probabilities of the four different trials. V0 refers to the absence of a visual stimulus. (**C**). Trial structure implied fixation cross presentation for 100 ms, followed 500 ms later by the presentation of one of two equally probable auditory stimuli for 600 ms, this, in turn, being followed 50 ms after its cessation by the presentation of one of two Gabor patches for 500 ms. Four target stimuli were also interleaved in a block after the ITI period. The representation of the two trials' time course shows that the target stimulus (task-relevant, lower line) was not always preceded by the audio–visual association (task-irrelevant, upper line). (**D**). Graphical description of the Hierarchical Gaussian Filter (HGF) model adopted to depict individual trajectories of precision-weighted prediction error on the basis of ongoing variability in weighting between sensory evidence and perceptual beliefs. (**E**). Exemplary single subject precision-weighted prediction error trajectory in relation to V1, V0 | A1 condition. (**F**). The EEG montage we adopted shows the position of the 27 channels.

2. Methods

2.1. Participants

Twenty-one volunteers (13 females, range 19–32, mean age 24.3 ± 3.4 (SD)) participated in the study. All were right-handed with normal or corrected-to-normal vision and normal hearing, they had no history of neurological disorders, and they were not taking any neurological medications. All participants gave informed written consent. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the University of Trento.

2.2. Procedure

During the experimental procedure participants were exposed to a stream of auditory and visual stimuli while sitting in a dimly lit booth at a distance of 1 m from the monitor (22.5″ VIEWPixx; resolution: 1024 × 768 pixels; refresh rate: 100 Hz; screen width: 50 cm). Participants were informed that they were involved in an audio–visual detection task consisting of a button response to the target stimuli presentation only. Auditory stimuli A1 and A2 consisted of low and high-frequency pure tones, 250 Hz and 500 Hz, respectively, whereas visual stimuli V1 and V2 consisted of two white-colored Gabor patches, sinusoidal gratings with a Gaussian envelope, with 45° and 135° orientation (4.4° × 3.4° visual angle, generated with Gaussian envelope, standard deviation = 18.0, spatial frequency = 0.08 cycles/pixel), presented against a grey background (Figure 1A). In each trial, auditory stimuli were followed by the presentation of visual stimuli according to an equivalent temporal sequence with two opposite probability distributions resulting in high frequent and low frequent visual stimulus occurrence (Figure 1B). In the condition of highly frequent visual stimulus occurrence, A1 stimulus was followed by V1 Gabor patch (V1 | A1) 90% of the time and 10% by visual stimulus absence, V0 | A1. In the condition of low frequent visual stimulus occurrence, the A2 stimulus was followed by a visual stimulus (V2 | A2) 10% of the time and by visual stimulus absence 90% of the time, V0 | A2. Pairings of the auditory and visual stimuli were counterbalanced across participants. Each trial started with a fixation cross presented for 100 ms, followed 500 ms later by the presentation of one of two equally probable auditory stimuli for 600 ms, and that, in turn, was followed 50 ms after its cessation by the presentation of one of two Gabor patches for 500 ms (Figure 1C). Trials of V0 | A1 and V0 | A2 conditions, which were of equal length, entailed no Gabor patch presentation. Trials were interspersed with an inter-trial interval (ITI) of 2500 ms ± 500 ms. The main audio–visual target detection task consisted of a button press response only when presented with specific target stimuli represented by an auditory target combining both the A1 and A2 stimuli, and a visual target combining both the V1 and V2 stimuli. Both of these lasted 500 ms and were followed by equal ITI. The experimental session consisted of 400 trials presented in 10 blocks, each including the random presentation of 4 perceptual targets, with a total duration of about 40 min. The experimental protocol was implemented using OpenSesame (v. 2.8) and PsychoPy (v. 2.1) as backend [62]. The experimental procedure aimed at inducing more attentional resource allocation to task-relevant perceptual targets. Importantly, probabilistic contingencies of audio–visual pairing were completely irrelevant to the audio–visual target detection task.

2.3. EEG Data Acquisition and Preprocessing

EEG data were recorded with a standard 10–5 system and 27 Ag/AgCl electrodes cap (EasyCap, Brain Products, Gilching, Germany) at a sampling rate of 1 kHz. Impedance was kept below 10 kΩ for all channels. AFz was used as the ground and the right mastoid was used as a reference. Electrodes were approximately evenly spaced and positioned at the following scalp sites: Fpz, Fz, F3, F4, F7, F8, F9, F10, FC5, FC6, T7, C3, Cz, C4, T8, CP5, CP6, P7, P3, Pz, P4, P8, PO7, PO8, O1, Oz, and O2 (Figure 1F). All preprocessing steps were conducted using EEGLAB [63] in accordance with guidelines and recommendations for EEG data preprocessing such as HAPPE [64], which are also applicable in the case of low-density recordings, HAPPILEE [65]. Spherical interpolation was carried out on a limited number of bad channels on the basis of channel correlation lower than 0.85 on average with respect to its neighbors and guided by visual inspection (average number of interpolated channels: 0.74, range: 0–3). Data were down-sampled at 250 Hz, high-pass filtered at 0.1 Hz and low-pass filtered at 80 Hz, using a Butterworth IIR filter with model order 2. CleanLine (https://github.com/sccn/cleanline accessed on 23 July 2023) with default parameters was used to remove power line 50 Hz noise and its harmonics up to 200 Hz. The data were then re-referenced to a common average reference [65] and epoched between −300 ms and 1300 ms relative to the onset of the auditory stimulus with a baseline correction between −300 ms and 0 ms. Artifact rejection was performed

through visual inspection and by an automatic procedure excluding epochs with very large signal amplitudes (detection threshold = ±500). The average number of trials rejected per participant was 1.1% (SD = 2.1%, range 0–7.3%). Stereotypical artefacts, including eyeblinks, eye movements, and muscle artefacts, were detected via independent component analysis using the extended Infomax algorithm [66]. A rejection strategy based on ICLabel [67] and visual inspection resulted in the removal of an average number of independent components equal to 9.33 (±3.48 SD). Finally, the data were converted to Fieldtrip format [68] for subsequent analyses.

2.4. EEG Data Analysis

Data analysis aimed at assessing implicit associative learning of paired audio–visual stimuli by investigating neural responses evoked by both predictive auditory stimuli and predicted visual stimulus presentation or omission. The analysis focused on two main epochs by performing grand averaging, considering auditory and visual stimuli separately: 0–650 ms relative to auditory stimuli presentation (audio period), and 650–1300 ms relative to Gabor patch presence or absence (post-audio period). Trials were divided into three equivalent groups to analyze the initial (first third of trials) and final phases (last third of trials) of sensory conditioning, and each phase consisted of 54 trials (except possible rejection of trials due to artifacts). Conventional event-related potential (ERP) analysis was first performed on both audio and post-audio periods considering initial and final trials. All trials of A1, A2, V1 | A1, and V0 | A2 conditions were averaged separately in relation to initial and final phases, whereas trials of V0 | A1 and V2 | A2 conditions were not analyzed due to their limited number required by the specific contingencies schedule of our conditioning paradigm. As in previous related studies [51,69], for ERP analysis we adopted an approach that considers the average neural response over comparable predefined regions of interest (ROI) using frontal (Fpz, Fz, F3, F4, F7, F8, F9, F10), temporo-parietal (FC5, FC6, T7, C3, Cz, C4, T8, CP5, CP6, P7, P3, Pz), and parieto-occipital channels (P4, P8, PO7, PO8, O1, Oz, O2), and thus enables more robust statistics through cluster-based correction. In addition, multivariate pattern classification based on Linear Discriminant Analysis (LDA) examining EEG signal differences between A1 and A2 at the subject level was performed considering all conditioning trials and channels as samples and features, respectively, with the MVPA-Light toolbox [70] and custom MATLAB (Mathworks, Natick, MA, USA) scripts. Z-scoring was applied across samples for each time point separately to normalize channel variances and remove baseline shifts. A 5-fold cross validation scheme was adopted and the Area Under the Curve (AUC) was used as a performance measure of LDA. An empirical chance level was obtained by running twice the same classification analysis with the same hyperparameter but with permuted labels. Statistical significance of model performance with respect to empirical chance level was assessed at the group level (paired permutation *t*-test two tailed, $\alpha = 0.05$) using mass univariate cluster-based permutation tests (10,000 iterations) and maxsum as cluster statistic, a valid and powerful way of dealing with the problem of multiple comparisons [71,72]. Effect size was estimated using Cohen's d (*d*) and Scaled Jeffrey-Zellner-Siow Bayes Factor *t*-test (BF_{10}, Cauchy prior with a scale parameter equal to $\sqrt{2}/2$), reporting the peak value inside the significant cluster. Analysis of relevant channels for classification was performed by converting the estimated weights of the LDA model at each fold into interpretable activation patterns [73]. Furthermore, we performed pattern similarity analysis to estimate the similarity of brain responses evoked by predictive A1 and A2 stimuli with those evoked by expected V1 and V0, by calculating Pattern Dissimilarity Matrices based on the cross-validated Mahalanobis distance (cvMD) [58] for each participant. This measure consists of splitting the trials in train and test sets, subtracting the average multivariate pattern across channels between conditions in both train and test sets, and finally performing a matrix multiplication between, the difference vector from the train set, the inverse of the covariance matrix estimated on the train set and the difference vector from the test set [58]. We computed the cvMD dissimilarity measure across trials between the auditory

period and the visual period at each time-point of the selected time window and at each ROI. A 5-fold cross validation scheme was applied to assess the similarity of EEG signal between audio (200–650 ms) and post-audio (0–550 ms) periods, considering time points (EEG samples) of initial and final trials. These time windows were selected as a result of the ERP analysis and MVPA. Statistical significance of Dissimilarity Matrices (DM) differences between initial and final trials was tested at group level with cluster-based permutation tests with the same hyperparameters as described above.

2.5. Hierarchical Gaussian Filter Modelling

The HGF is a Bayesian generative model [61,74] of perceptual inference on a changing environment based on sequential input [13,75]. The HGF consists of perceptual and response models, representing a Bayesian ideal observer who receives a sequence of inputs and generates behavioral responses. Since our experimental design deliberately precluded behavioral responses, we used only the perceptual model [76]. In this framework, a perceptual model comprised 3 hierarchal hidden states (x), which accounted for a multi-level belief updating process of the hierarchically related environmental states giving rise to sensory inputs, and the observed input (y) representing the actual occurrence of a stimulus in a given trial (Figure 1D).

Our HGF model assumed that environmental hidden states evolved conditionally on the states at the immediately higher level. The hidden states processed at the first level of the perceptual model represented a sequence of beliefs (x_1^t) about stimulus occurrence, that is, whether a visual stimulus was presented ($y^t = 1$) or absent ($y^t = 0$) at trial t, and was modelled as follows:

$$x_1^t \left| x_2^t \right. \sim Bernoulli\left(s\left(x_2^t\right)\right), \tag{1}$$

where $s\left(x_2^t\right) := \left[1 + exp\left(x_2^t\right)\right]^{-1}$ is the logistic sigmoid function. Here, the hidden states at the second level (x_2^t) is an unbounded real parameter of the probability that $x_1^t = 1$, thus representing the current belief of the probability that a given stimulus occurred. Such a hidden state process evolves according to a Gaussian random walk:

$$x_2^t \left| x_2^{t-1}, x_3^t \right. \sim Gaussian\left(x_2^{t-1}, exp\left(\kappa x_3^t + \omega\right)\right), \tag{2}$$

which depends on both its value at a previous trial t, and the hidden state at the third level of the hierarchy. In particular, the higher-level hidden state process ($x_3^{(t)}$) determines the log-volatility of the hidden state process at the second level, thus codifying the volatility of the environment during the time course of the experiment. This process evolves according to a Gaussian random walk:

$$x_3^t \left| x_3^{t-1} \right. \sim Gaussian\left(x_3^{t-1}, \vartheta\right). \tag{3}$$

The parameter set ($\kappa, \omega, \vartheta$) determined the dispersion of the random walks at different levels of the hierarchy and allowed us to shape individual difference in learning. By inverting the generative model, given a sequence of observations (y), it was possible to obtain the updating process of the trial-by-trial estimates of the hidden state variables.

The update rules shared a common structure across the model's hierarchy: at any level i the update of the posterior mean μ_i^t of the state x_i, that represented the belief on trial k, was proportional to the precision-weighted prediction error (pwPE) ε_i^t as follows:

$$\mu_i^{t-1} - \mu_i^t \propto \psi_i^t \delta_{i-1}^t = \varepsilon_i^t, \tag{4}$$

$$\psi_i^t = \frac{\hat{\pi}_{i-1}^t}{\pi_i^t}, \tag{5}$$

$$\pi_i^t = \frac{1}{\sigma_i^t}. \tag{6}$$

As shown in Equations (4)–(6), in each trial, a belief update $\mu_i^{t-1} - \mu_i^t$ is proportional to the prediction error at the level below δ_{i-1}^t. The pwPE is the product of the prediction error δ_{i-1}^t and a precision ratio ψ_i^t that depends on the precision (inverse variance, Equation (5)) of the prediction at the level below $\hat{\pi}_{i-1}^t$ and the current level π_i^t. In this application, we were interested in the updated equations of the hidden states at the second level, which have a general form similar to those of traditional reinforcement learning models, such as the Rescorla-Wagner model [77]. The pwPE on the second level, was thus assumed to be responsible for the learned perceptual associations. The nature of the pwPE could be described through the following update equation of the mean of the second level:

$$\mu_2^t = \mu_2^{t-1} + \sigma_2^t\left(\mu_1^t - s\left(\mu_2^{t-1}\right)\right), \tag{7}$$

where the last term represents the prediction error $\left(\mu_1^t - s\left(\mu_2^{t-1}\right)\right)$ at the first level weighted by the precision term σ_2^t (see [61] for a general derivation and more mathematical details). Individual trajectories of pwPEs with separate models for V1, V0 | A1 and V2, V0 | A2 were calculated by estimating the parameters that minimized Bayesian Surprise using the Broyden–Fletcher–Goldfar–Shannon (BFGS) quasi-Newton optimization algorithm. We determined these Bayes optimal perceptual parameters by inverting the perceptual model based on the stimulus sequence alone and a predefined prior for each parameter (HGF toolbox, version 5.2 implemented via the Translational Algorithms for Psychiatry Advancing Science toolbox). These model-derived trajectories of pwPEs (Figure 1E) from the second level were used as regressors in a general linear model (GLM) applied to each channel-time point pair for each participant. We used the R^2 measure for evaluating the goodness of fit and averaged these values over the selected ROIs. Statistical significance was tested at the group level with cluster-based permutation tests and using the same hyperparameters as previously described.

3. Results

We collected EEG and behavioral data from 21 human volunteers exposed to a stream of non-target auditory and visual stimuli, whilst being involved in a main audio–visual detection task consisting of button press responses to target stimuli presentation (Figure 1A). Crucially, we manipulated the transition probabilities between the non-target stimuli so that non-target audio–visual co-occurrences had no predictable effects on target presentation, thus making the learning of these associations task-irrelevant (Figure 1B,C). Participants debriefed at the end of the experiment reported not being consciously aware of stimuli pairings. When specifically questioned about possible stimuli associations they reported not to have noticed either any regularities of stimuli presentation or specific audio–visual pairings.

3.1. Event-Related Potentials

We performed conventional event-related potential (ERP) analysis on both audio and post-audio periods considering the initial and final trials so as to assess the evoked activity throughout the learning of statistical associations of non-target auditory and visual stimuli. Trials were divided into three equivalent groups to analyze the initial (first third of trials) and final phases (last third of trials) of sensory conditioning. Results of ERP analysis comparing initial versus final trials of sensory conditioning in relation to audio period revealed a significant attenuation of signal amplitude in response to A1 auditory stimulus, predictive of V1 visual stimulus (V1 | A1), in the interval 190–280 ms in the occipital ROI (Figure 2A, $d = 0.92$, $BF_{10} = 78.31$, $p = 0.0045$, cluster corrected) and in response to A2 auditory stimulus, predictive of V0 stimulus (stimulus absence), in the interval 180–240 in the temporo-parietal ROI (Figure 2A, $d = 0.79$, $BF_{10} = 22.97$, $p = 0.0278$, cluster corrected). Considering only final trials, a reduced negativity was observed for A1 with respect to A2 in the interval 180–230 ms in the frontal ROI (Figure 2A, $d = 0.77$, $BF_{10} = 19.57$, $p = 0.0260$, cluster corrected) and in the interval 180–285 ms in the temporo-parietal ROI (Figure 2A,

$d = 1.02$, $BF_{10} = 211.36$, $p = 0.008$, cluster corrected). In the post-audio period, comparison of initial and final trials revealed a significant attenuation of the response to V1 | A1 in the interval 60–170 ms in the frontal ROI (Figure 2B, $d = 0.84$, $BF_{10} = 38.03$, $p = 0.0155$, cluster corrected) and in the intervals −10–160 ms (Figure 2B, $d = 1.11$, $BF_{10} = 443.03$, $p = 0.0134$, cluster corrected) and 530–600 ms (Figure 2B, $d = 0.64$, $BF_{10} = 5.82$, $p = 0.0236$, cluster corrected) in the parieto-occipital ROI. A significant signal attenuation was also observed for V0 | A2 condition in the interval −50–40 ms in the frontal ROI (Figure 2B, $d = 0.69$, $BF_{10} = 9.25$, $p = 0.0466$, cluster corrected) and in the interval −30–30 ms in the parieto-occipital ROI (Figure 2B, $d = 0.75$, $BF_{10} = 15.95$, $p = 0.0321$, cluster corrected). In short, these results revealed across conditioning changes in neural responses to both predictive auditory and predicted visual stimuli, that is the acquired conditioned response and unconditioned response, respectively.

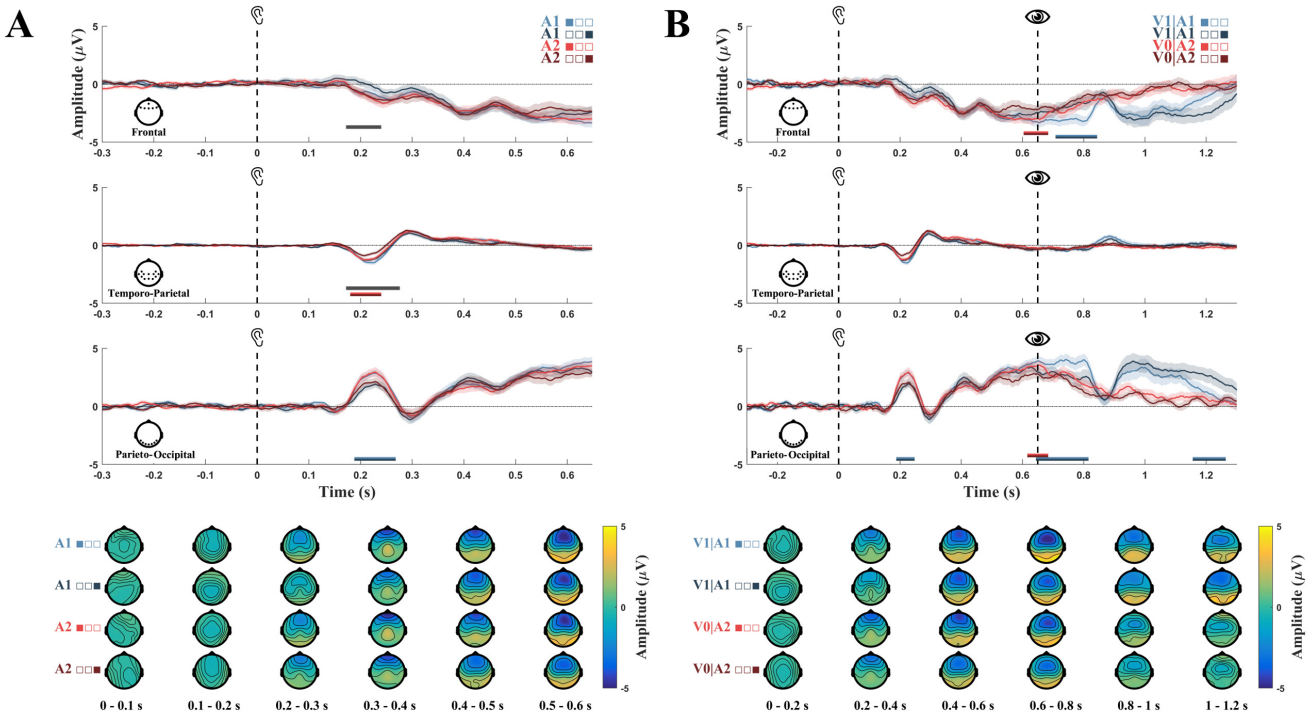


Figure 2. (**A**). Top, ERPs of initial (A1 light blue, A2 light red) and final (A1 dark blue, A2 dark red) trials in relation to auditory stimuli onset for each selected ROI. Shading refers to SEM across participants, horizontal bars refer to statistical significance. Bottom, topographic maps depicting whole-brain spatial distribution of EEG signal across 100 ms intervals after auditory stimuli onset. In the legend, filled squares indicate the initial or final blocks of trials from which the ERPs are computed (**B**). Top, Full time-course ERPs of initial (V1 | A1 light blue, V0 | A2 light red) and final (V1 | A1 dark blue, V0 | A2 dark red) trials including both auditory and visual stimuli presentation for each selected ROI. Bottom, Topographic maps depicting whole-brain spatial distribution of EEG signal across 100 ms intervals after auditory stimuli onset.

3.2. Multivariate Classification

Multivariate decoding, performed using Linear Discriminant Analysis-based classification, aimed to assess EEG signal differences between A1 and A2 during the whole conditioning session. Such expected differences should reflect differential processing of equivalent auditory stimuli that distinctively anticipated visual stimulus presence or absence. Our results showed significant discrimination performance between A1 and A2 with respect to chance level in three different time windows (Figure 3A): 192–340 ms ($AUC_{peak} = 0.56$, $d = 1.17$, $BF_{10} = 771.98$, $p < 0.0001$, cluster corrected), 344–444 ms ($AUC_{peak} = 0.53$, $d = 0.85$, $BF_{10} = 39.42$, $p = 0.0017$, cluster corrected), and 500–540 ms

($AUC_{peak} = 0.52$, $d = 0.73$, $BF_{10} = 13.04$, $p = 0.0484$, cluster corrected), that is immediately preceding probabilistic visual stimulus presence or absence. These results thus indicated distinct processing of equivalent A1 and A2 stimuli, differentially predicting V1 and V0, respectively. Relevant channels for classification were located in the frontal regions for the first significant time window and mostly in the temporo-occipital regions for the other two time windows.

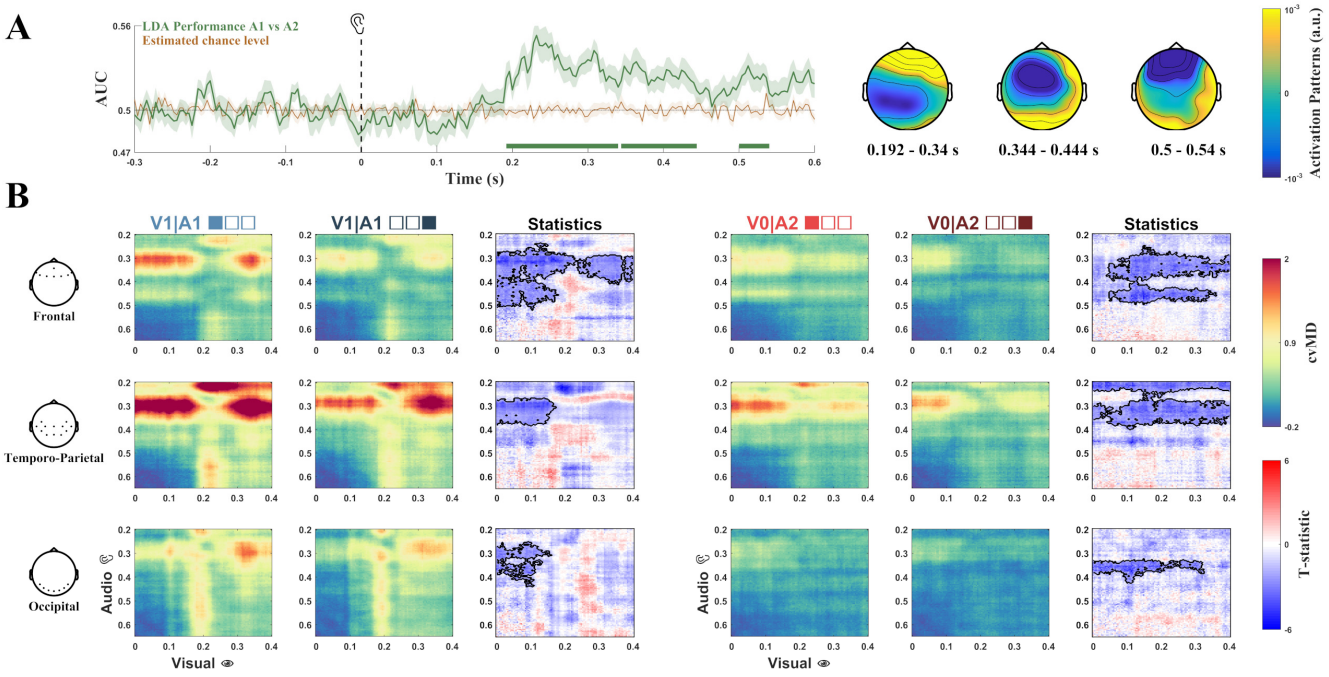


Figure 3. (**A**). Left, EEG signals classification performance (Area Under the Curve) of A1 versus A2 during audio period (green) against estimated chance level (orange). Shading indicates SEM across folds. Green horizontal lines indicate statistical significance. Right, topographic maps of activation patterns computed from the model's parameters, representing feature importance for classification performance in relation to EEG channels and time windows. (**B**). Pattern Similarity Analysis. PSA aimed to test whether the pattern elicited by the predictive auditory stimulus increasingly resembled the pattern elicited by the expected visual stimulus during sensory conditioning. The images show Pattern Dissimilarity Matrices of EEG signal considering audio and post-audio periods for V1 | A1 (left) and V0 | A2 (right) condition, for each selected ROI. For each condition, the first column refers to cvMD during initial trials, the second column refers to cvMD during final trials, and the third column depicts statistical differences between initial and final trials. Outlined regions indicate statistical significance. X and Y axes depict the time course of visual and auditory stimulus, respectively. For the visual stimulus, the temporal interval ranges between visual stimulus onset and its offset, whereas for the auditory stimulus, it ranges between 200 ms after the auditory stimulus onset (this time was set on the basis of ERP and MVPA classification results) and its offset.

3.3. Pattern Similarity Analysis

Pattern similarity analysis aimed to test whether the pattern elicited by the predictive auditory stimulus increasingly resembled the pattern elicited by the expected visual stimulus during sensory conditioning. Pattern Similarity Analysis (PSA) was performed by calculating a dissimilarity matrix (DM) between audio and post-audio periods to specifically assess the relationship of brain responses evoked by predictive A1 and A2 stimuli with those evoked by predicted V1 and V0, respectively (Figure 3B). DM, estimated on the basis of the cross-validated Mahalanobis distance (cvMD), showed that in the V1 | A1 condition the cvMD was significantly lower in the final trials with respect to initial trials in several clusters in frontal, temporo-parietal, and occipital ROIs. In the frontal ROI two main significant clusters were observed, one corresponding to ~300–500 ms in the audio

period and ~0–200 ms in the post-audio period ($d = 1.36$, $BF_{10} = 4608.53$, $p < 0.0001$, cluster corrected), and another one corresponding to ~300–400 ms in the audio period and ~250–400 ms in the post-audio period ($d = 0.84$, $BF_{10} = 36.18$, $p = 0.0009$, cluster corrected). In the temporo-parietal ($d = 1.02$, $BF_{10} = 202.98$, $p = 0.005$, cluster corrected) and parieto-occipital ($d = 0.99$, $BF_{10} = 161.58$, $p = 0.015$, cluster corrected), ROIs the significant cluster corresponded to ~300–400 ms in the audio period and ~0–150 ms in the post-audio period. In V0 | A2 condition, the cvMD was also significantly lower in the final trials with respect to initial trials in several clusters in frontal, temporo-parietal, and occipital ROIs. In the frontal ROI two main significant clusters were observed, one corresponding to ~300–400 ms in the audio period and ~50–400 ms in the post-audio period ($d = 1.07$, $BF_{10} = 313.13$, $p = 0.0048$, cluster corrected), and another one corresponding to ~400–500 ms in the audio period and ~50–350 ms in the post-audio period ($d = 1.21$, $BF_{10} = 1132.09$, $p = 0.005$, cluster corrected). In the temporo-parietal ROI two significant clusters were also observed, one corresponding to ~200–250 ms in the audio period and ~0–400 ms in the post-audio period ($d = 1.10$, $BF_{10} = 434.36$, $p = 0.002$, cluster corrected), and another one corresponding to ~280–380 ms in the audio period and ~0–500 ms in the post-audio period ($d = 1.03$, $BF_{10} = 219.11$, $p = 0.0019$, cluster corrected). In addition, a significant cluster was observed in the occipital ROI corresponding to ~320–400 ms in the audio period and ~0–300 ms in the post-audio period ($d = 1.21$, $BF_{10} = 1124.35$, $p = 0.007$, cluster corrected). Altogether these results showed decreased dissimilarity for both V1 | A1, an V0 | A2 conditions during sensory conditioning.

3.4. HGF Modelling

We modeled individual trajectories of precision weighted prediction error (pwPE) on the basis of ongoing variability in weighting between sensory evidence and perceptual beliefs using the Hierarchical Gaussian filter (HGF), a Bayesian ideal observer model that attempts to predict future stimuli occurrence on the basis of the history and uncertainty of contextual events. This analysis allowed us to assess differential prediction error processing associated with violation of expected presentation or omission of visual stimulus, and to further demonstrate that associative learning of perceptual stimuli can occur even when perceptual stimuli are tasirrelevant. We thus fed into the HGF model the same sequence of stimuli that every subject was exposed to, and then we obtained the prediction error trajectory. We used these ideal error trajectories as regressors in a general linear model (GLM) applied to each channel-time point pair, for each participant. HGF analysis showed that pwPE, significant for both V1, V0 | A1 and V2, V0 | A2, was mediated by all selected brain regions, although the largest effect size was in the occipital ROI (Figure 4). Regression analysis resulted in a significant effect for both V1, V0 | A1 and V2, V0 | A2 (Figure 4) in the interval ~250–550 ms of post-audio period in the frontal (V1, V0 | A1: $d = 0.66$, $BF_{10} = 7.29$, $p = 0.0007$, cluster corrected; V2, V0 | A2: $d = 0.72$, $BF_{10} = 11.84$, $p = 0.0007$, cluster corrected) and temporo-parietal ROIs (V1, V0 | A1: $d = 0.81$, $BF_{10} = 26.01$, $p = 0.0003$, cluster corrected; V2, V0 | A2: $d = 0.69$, $BF_{10} = 8.89$, $p = 0.0005$, cluster corrected), and ~150–550 ms in the occipital ROI (V1, V0 | A1: $d = 0.72$, $BF_{10} = 11.96$, $p = 0.0386$, cluster corrected; V2, V0 | A2: $d = 0.77$, $BF_{10} = 18.89$, $p < 0.0001$, cluster corrected). Finally, R^2 was significantly larger for V1, V0 | A1 than for V2, V0 | A2 (Figure 4) in the interval ~370–400 ms in the post-audio period in the occipital ROI ($d = 0.64$, $BF_{10} = 6.03$, $p = 0.0398$, cluster corrected), indicating differential neural processing of PE in relation to violation of predicted visual stimulus presentation and omission.

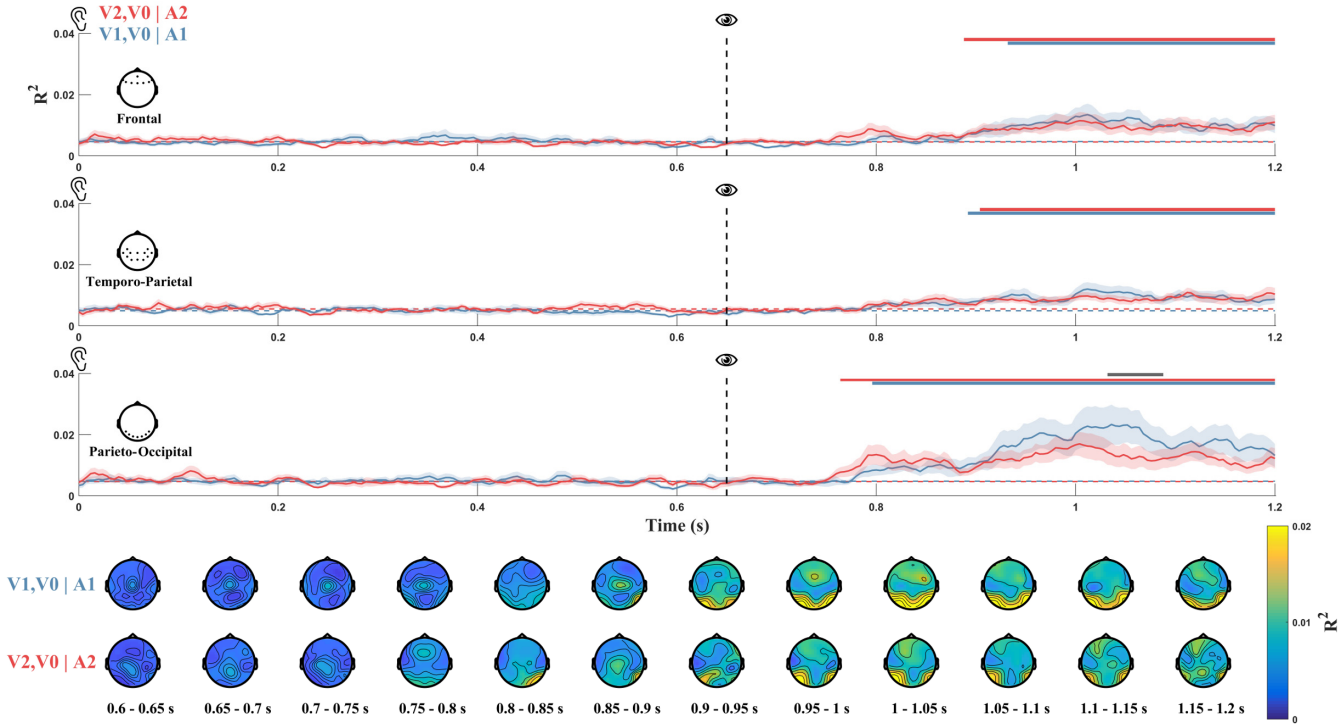


Figure 4. (**Top**). Line plots of GLM performance (R^2) fitting HGF model-derived precision-weighted prediction error trajectories to each EEG channel, averaged across ROIs, for V1, V0 | A1 (solid blue) and V2, V0 | A2 (solid red) conditions. Dotted lines represent average performances during the baseline period (preceding the onset of visual stimuli). Shading indicates SEM across participants. Horizontal lines indicate the statistical significance of differences between each model and baseline (blue and red for V1, V0 | A1 and V2, V0 | A2, respectively), and between V1, V0 | A1 and V2, V0 | A2 (grey). (**Bottom**). Topographic maps depicting whole-brain spatial distribution of GLM performance for each model across 50 ms intervals in the post-audio period.

4. Discussion

In our study, we investigated whether incidental but recurrent exposure to non-salient and task-irrelevant probabilistic audio–visual associations induced implicit learning of perceptual patterns resulting in conditioned cortical responses, possibly representing early instantiations of predicted stimulus-related sensory information. To this aim, we assessed both temporal and representational aspects of neural signals associated with the learned conditional probability of paired audio–visual stimuli. Our analyses showed that incidental perceptual regularities were rapidly encoded, leading to differential neural responses to anticipatory auditory stimuli predictive of either the presence or absence of a visual stimulus. ERPs results revealed that in the final trials of sensorial conditioning brain responses to equivalent A1 and A2 stimuli, anticipating probabilistic visual stimulus presence and absence, respectively, were significantly different around 200 ms after stimulus onset, in both the frontal and temporo-parietal regions (Figure 2A,B). Specifically, a response attenuation in the temporo-parietal region was observed in response to A2 but not to A1. This suggests that A1, evoking anticipation of V1 being presented in contrast to A2 preceding an absence of visual stimulus, required differential processing.

In line with the known audiovisual cross-modal effect resulting in visual cortex activity during auditory perception [78,79], increased signal amplitude in the parieto-occipital region was observed in response to both A1 and A2 across all trials. However, a significant attenuation over time in the parieto-occipital region was measured in response to A1, but not to A2. The point of repetition suppression effect only being observed for A1 indicates differential predictive processes in relation to expected visual stimulus presentation and omission.

According to the assumption that the expectation suppression effect reflects decreased surprise signals, we would have expected to observe a similar attenuation for both A1 and A2. Vice versa, in line with the interpretation of sensory attenuation resultant of the sharpening of stimulus-specific representation [2,9,28], we interpreted the parieto-occipital differential response to A1 with respect to A2 as increased response tuning. This is because A1, in contrast to A2, carries information about consecutive V1 presentations. Our results are also in line with previous evidence of enhanced stimulus-specific baseline activity during early sensory processing in relation to probabilistic signal occurrence [14]. Multivariate pattern analysis corroborated ERP results showing differential neural responses to A1 and A2, in particular in a temporal interval immediately preceding probabilistic visual stimulus presence or absence and disclosed that such divergent responses manifested early in the frontal region and later in the temporo-occipital areas (Figure 3A). These results might also be potentially ascribable to the inherent difference of auditory stimuli. However, the analysis of feature relevance, showing that occipital channels were particularly involved in later stages of the classification (350–600 ms), suggests that the two auditory stimuli might indeed convey differential predictive information of consequent stimulus presentation. Pattern similarity analysis revealed that for the V1 | A1 condition the cvMD between audio and post-audio period, although initially large, decreased over time in the frontal, parieto-occipital, and temporo-parietal regions, resulting in decreased dissimilarity of neural activity between ~0–150 ms after V1 onset and ~250–400 ms after A1 onset (Figure 3B). Decreased cvMD was also observed in the frontal region between neural activity ~400–500 ms after A1 onset and ~0–150 ms after V1 onset (Figure 3B).

Altogether, our findings indicated that over time the pattern of neural activity elicited by A1 increasingly resembled that elicited by V1, and thus suggest that some kind of stimulus-specific perceptual priors relative to probabilistic but task-irrelevant audio–visual association were instantiated at the cortical level during the unfolding of sensory conditioning. Previous studies showed that prior expectations can elicit anticipatory sensory-like neural representations of predictable sensory information, in particular when this is somehow related to the task [28,32]. In particular, Kok and colleagues reported that a predictive auditory cue evoked early sensory representation in the primary visual cortex immediately before the expected visual stimulus [32], or at the time of the expected but omitted visual stimulus [28]. In our study, in line with previous findings showing an expectation-induced decrease of primary visual cortex activity [9,10,28,80], we also measured an attenuated response over time during V1 | A1, and immediately preceding V1, in the parieto-occipital and frontal regions. Similarity analysis did not show evidence of pre-stimulus neural instantiation of expected sensory input immediately before visual stimulus presentation (−50 to 0 ms) as in Kok et al. [32] but a similar effect occurred earlier at the presentation of the predictive auditory stimulus.

Notably, the cvMD also decreased over time in relation to V0 | A2 condition in the frontal, parieto-occipital, and temporo-parietal regions. In particular, decreased A2-related cvMD in the temporo-parietal region was observed between neural activity around 200–250 ms and 280–350 ms in the audio period, and in the V0–V2 post-audio period (0–400 ms). These results are possibly ascribable to the attenuated response to A2, resulting in increased similarity with the response to frequent V2 omission. On the other hand, as for the A1 stimulus, these effects might also reflect the tuning of A2 response progressively incorporating perceptual priors of V2 occurrence's low probability. Accordingly, decreased dissimilarity was also observed in the parieto-occipital region between neural activity occurring 320–400 ms after A2 and during the V0–V2 time period. Moreover, neural attenuation over time was also observed for the V0 | A2 condition, in both frontal and parieto-occipital regions, in a time window anticipating and also corresponding to V2 presentation. This further suggests an increased encoding of low V2 occurrence probability.

Overall, our results indicated that forward neural processes anticipating some specific aspects of consecutive visual stimulus presentation might occur earlier during the initial

processing of predictive stimulus; however, due to limited spatial resolution, the exact representational content of such anticipatory activity has yet to be fully clarified.

As reported in previous studies [10,21], we observed anticipatory neural mechanisms independently of stimulus salience, as our stimuli consisted of non-salient abstracts and auditory and visual stimuli. In addition, the observed effects were task-independent, since statistical regularity of paired audiovisual stimuli was not necessary and irrelevant to detecting the target stimuli. This result is also in line with studies showing that task-irrelevant visual perceptual learning can occur as a result of mere exposure to perceptual features.

It has been proposed that several aspects of perceptual and statistical learning might be unified in the framework of Hierarchical Bayesian [81]. In both perceptual and statistical learning, attention plays an important role [45,82]. In visual perceptual learning, attention enhances bottom-up signals from task-relevant features, whereas it decreases signals from task-irrelevant features; however, visual perceptual learning of task-irrelevant features can also occur as far as they can be optimally expected (suprathreshold presentation) [82]. Similarly, in statistical learning, attenuation of neural responses to predictable stimuli vanishes when they are not expected [45]. In our study, as non-target stimuli shared both auditory and visual features with target stimuli, we supposed that they were likely expected, however overt encoding of the statistical regularity of audio–visual associations was not necessarily completely irrelevant to the task.

Finally, the observed conditioned responses in the sensory cortices appeared to be mediated by frontal and prefrontal areas. The differential responses to A1 and A2 in the frontal ROI, preceding changes in temporo-occipital ROIs suggest that prefrontal regions might support mutual information exchange between auditory and visual cortices [83], likely in relation to temporal aspects of perceptual regularities [84–86] for timely instantiation of specific perceptual priors. Sensory priors inducing auditory-cued shaping of visual cortex responses in both V1 | A1 and V0 | A2 conditions might then be mediated by direct interactions between auditory and visual cortices [80].

Remarkably, computational modeling of pwPE trajectories further demonstrated the instantiation of predictive neural signals by showing distinct neural processing of prediction error in relation to violations of expected visual stimulus presence or absence. Moreover, differential neural processing of PE correlated with activity in frontal, temporal, and occipital areas (Figure 4) at latencies corresponding to those of typical event-related potentials elicited by deviant stimuli [87,88]. The precision of predictions for audiovisual patterns, likely mediated by temporal regions such as the superior temporal gyrus [89,90], might trigger a gradual update of prefrontal regions-mediated cortical representations of expected V1 and V0 [49,91]. Indeed, analysis of pwPE trajectories revealed a significant difference of V1 | A1 with respect to V0 | A2 occurring about 300–400 ms after V1 and V0 onset in the occipital region, indicating stimulus-specific differential processing of prediction violation [92,93].

5. Conclusions

Our study demonstrated rapid encoding of incidental but probabilistic audio–visual regularities leading to stimulus-specific modulation of cortical responses to both predictive and predicted sensory stimuli. These findings corroborate previous evidence of statistical learning of task-irrelevant sensory regularities [21,29] and extend it by showing that the acquired sensory associations lead to specific predictive processes resembling those observed in relation to learned task-relevant sensory patterns. In particular, we showed that sensory conditioning of task-irrelevant audio–visual associations appeared to induce increased similarity between the neural response to predictive auditory stimuli and the response to predicted visual stimulus presence or absence. Remarkably, Hierarchical Gaussian filter modeling, estimating dynamic changes of prediction error signals in relation to differential probabilistic occurrences of audio–visual stimuli, further demonstrated instantiation of predictive neural signals by showing distinct neural processing of prediction error in relation to violations of expected visual stimulus presence or absence. Overall, our findings suggest

that statistical learning of non-salient and task-irrelevant perceptual regularities might induce early generation of neural priors at the time of predictive stimulus presentation conveying sensory-specific information of the predicted consecutive stimulus. However, the exact nature of sensory-specific priors is yet to be elucidated, in particular to what extent predictive stimuli anticipate the expected consecutive visual stimulus. Further studies using more conventional conditioning paradigms and using multiple predictable sensory stimuli with differential probabilistic schemes, might help to clarify this aspect. Moreover, advanced techniques such as high-resolution EEG and fMRI retinotopic mapping might definitely permit delineating a clearer representational pattern elicited by the predictive neural signals. Finally, it remains unclear whether encoding of statistical regularities of sensory stimuli occurred implicitly or explicitly. Explicit attention to task-irrelevant audio–visual stimuli was unnecessary to perform our target detection task, on the other hand, our target stimuli shared some features with either auditory or visual stimuli that possibly biased attention.

Nevertheless, here as in the case of goal-directed behavior where predictive mechanisms were related to information-seeking random exploration [91,94], the intrinsic neural need to reduce uncertainty about state transition dynamics of the environment might also explain learned incidental probabilistic sensory patterns. Learning sensory representations might then depend on automatic perceptual processes exploiting either reward statistics of past experience or beliefs about future representations [95] that optimize neural computations for adaptive behavior.

Ultimately, in accordance with predictive brain principles, our results suggest that associative processes might even occur exclusively at the perceptual level, possibly as a consequence of Hebbian neural plasticity [96], and that stimulus salience that is typically considered a critical element for learning in classical conditioning might not, in fact, be strictly required [97,98].

Author Contributions: A.G.: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Writing—original draft, and Writing—Review and Editing; M.D.: Software, Formal analysis, Writing—Review and Editing; G.G.: Investigation, Writing—Review and Editing; C.R.: Software, Formal analysis, Writing—Review and Editing; C.B.: Methodology, Writing-Review & Editing; A.C.: Conceptualization, Methodology, Writing- original draft, Writing—Review and Editing, Supervision, and Project administration. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a BIAL foundation grant to Andrea Caria (Grant No. 137/22).

Institutional Review Board Statement: The study was approved by the ethics committee at the University of Trento (protocol n° 2018-009) and conformed to the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets generated and analyzed during the current study are not publicly available since participants did not provide explicit written consent regarding the sharing of their data on public repositories, but are available from the corresponding author on reasonable request and the requirement for co-authorship or inclusion in the author byline.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Arnal, L.H.; Giraud, A.L. Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* **2012**, *16*, 390–398. [CrossRef] [PubMed]
2. de Lange, F.P.; Heilbron, M.; Kok, P. How Do Expectations Shape Perception? *Trends Cogn. Sci.* **2018**, *22*, 764–779. [CrossRef] [PubMed]
3. Friston, K. The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **2010**, *11*, 127–138. [CrossRef]
4. Friston, K. The free-energy principle: A rough guide to the brain? *Trends Cogn. Sci.* **2009**, *13*, 293–301. [CrossRef] [PubMed]
5. Press, C.; Kok, P.; Yon, D. The Perceptual Prediction Paradox. *Trends Cogn. Sci.* **2020**, *24*, 13–24. [CrossRef] [PubMed]
6. Wacongne, C.; Changeux, J.P.; Dehaene, S. A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* **2012**, *32*, 3665–3678. [CrossRef] [PubMed]

7. Bastos, A.M.; Lundqvist, M.; Waite, A.S.; Kopell, N.; Miller, E.K. Layer and rhythm specificity for predictive routing. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 31459–31469. [CrossRef] [PubMed]
8. Friston, K.J.; Stephan, K.E. Free-energy and the brain. *Synthese* **2007**, *159*, 417–458. [CrossRef] [PubMed]
9. Kok, P.; Jehee, J.F.; de Lange, F.P. Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron* **2012**, *75*, 265–270. [CrossRef] [PubMed]
10. Kok, P.; Failing, M.F.; de Lange, F.P. Prior expectations evoke stimulus templates in the primary visual cortex. *J. Cogn. Neurosci.* **2014**, *26*, 1546–1554. [CrossRef] [PubMed]
11. Malekshahi, R.; Seth, A.; Papanikolaou, A.; Mathews, Z.; Birbaumer, N.; Verschure, P.F.; Caria, A. Differential neural mechanisms for early and late prediction error detection. *Sci. Rep.* **2016**, *6*, 24350. [CrossRef] [PubMed]
12. Melloni, L.; Schwiedrzik, C.M.; Muller, N.; Rodriguez, E.; Singer, W. Expectations change the signatures and timing of electrophysiological correlates of perceptual awareness. *J. Neurosci.* **2011**, *31*, 1386–1396. [CrossRef]
13. Powers, A.R.; Mathys, C.; Corlett, P.R. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* **2017**, *357*, 596–600. [CrossRef] [PubMed]
14. Wyart, V.; Nobre, A.C.; Summerfield, C. Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 3593–3598. [CrossRef] [PubMed]
15. Sherman, B.E.; Graves, K.N.; Turk-Browne, N.B. The prevalence and importance of statistical learning in human cognition and behavior. *Curr. Opin. Behav. Sci.* **2020**, *32*, 15–20. [CrossRef] [PubMed]
16. Bar, M. The proactive brain: Using analogies and associations to generate predictions. *Trends Cogn. Sci.* **2007**, *11*, 280–289. [CrossRef] [PubMed]
17. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **2013**, *36*, 181–204. [CrossRef]
18. Enns, J.T.; Lleras, A. What's next? New evidence for prediction in human vision. *Trends Cogn. Sci.* **2008**, *12*, 327–333. [CrossRef] [PubMed]
19. Rao, R.P.; Ballard, D.H. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **1999**, *2*, 79–87. [CrossRef]
20. Alink, A.; Schwiedrzik, C.M.; Kohler, A.; Singer, W.; Muckli, L. Stimulus predictability reduces responses in primary visual cortex. *J. Neurosci.* **2010**, *30*, 2960–2966. [CrossRef] [PubMed]
21. den Ouden, H.E.; Friston, K.J.; Daw, N.D.; McIntosh, A.R.; Stephan, K.E. A dual role for prediction error in associative learning. *Cereb. Cortex* **2009**, *19*, 1175–1185. [CrossRef] [PubMed]
22. Summerfield, C.; Trittschuh, E.H.; Monti, J.M.; Mesulam, M.M.; Egner, T. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* **2008**, *11*, 1004–1006. [CrossRef] [PubMed]
23. Summerfield, C.; Wyart, V.; Johnen, V.M.; de Gardelle, V. Human Scalp Electroencephalography Reveals that Repetition Suppression Varies with Expectation. *Front. Hum. Neurosci.* **2011**, *5*, 67. [CrossRef] [PubMed]
24. Feuerriegel, D.; Vogels, R.; Kovacs, G. Evaluating the evidence for expectation suppression in the visual system. *Neurosci. Biobehav. Rev.* **2021**, *126*, 368–381. [CrossRef] [PubMed]
25. Walsh, K.S.; McGovern, D.P.; Clark, A.; O'Connell, R.G. Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Ann. N. Y. Acad. Sci.* **2020**, *1464*, 242–268. [CrossRef] [PubMed]
26. Garrido, M.I.; Rowe, E.G.; Halasz, V.; Mattingley, J.B. Bayesian Mapping Reveals That Attention Boosts Neural Responses to Predicted and Unpredicted Stimuli. *Cereb. Cortex* **2018**, *28*, 1771–1782. [CrossRef]
27. Wacongne, C.; Labyt, E.; van Wassenhove, V.; Bekinschtein, T.; Naccache, L.; Dehaene, S. Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20754–20759. [CrossRef] [PubMed]
28. Aitken, F.; Menelaou, G.; Warrington, O.; Koolschijn, R.S.; Corbin, N.; Callaghan, M.F.; Kok, P. Prior expectations evoke stimulus-specific activity in the deep layers of the primary visual cortex. *PLoS Biol.* **2020**, *18*, e3001023. [CrossRef] [PubMed]
29. McIntosh, A.R.; Cabeza, R.E.; Lobaugh, N.J. Analysis of neural interactions explains the activation of occipital cortex by an auditory stimulus. *J. Neurophysiol.* **1998**, *80*, 2790–2796. [CrossRef]
30. Meyer, T.; Olson, C.R. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19401–19406. [CrossRef]
31. Sakai, K.; Miyashita, Y. Neural organization for the long-term memory of paired associates. *Nature* **1991**, *354*, 152–155. [CrossRef] [PubMed]
32. Kok, P.; Mostert, P.; de Lange, F.P. Prior expectations induce prestimulus sensory templates. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 10473–10478. [CrossRef] [PubMed]
33. SanMiguel, I.; Widmann, A.; Bendixen, A.; Trujillo-Barreto, N.; Schroger, E. Hearing silences: Human auditory processing relies on preactivation of sound-specific brain activity patterns. *J. Neurosci.* **2013**, *33*, 8633–8639. [CrossRef] [PubMed]
34. Blom, T.; Feuerriegel, D.; Johnson, P.; Bode, S.; Hogendoorn, H. Predictions drive neural representations of visual events ahead of incoming sensory information. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 7510–7515. [CrossRef] [PubMed]
35. Boettcher, S.E.P.; Stokes, M.G.; Nobre, A.C.; van Ede, F. One Thing Leads to Another: Anticipating Visual Object Identity Based on Associative-Memory Templates. *J. Neurosci.* **2020**, *40*, 4010–4020. [CrossRef] [PubMed]
36. Zhou, Y.J.; Perez-Bellido, A.; Haegens, S.; de Lange, F.P. Perceptual Expectations Modulate Low-Frequency Activity: A Statistical Learning Magnetoencephalography Study. *J. Cogn. Neurosci.* **2020**, *32*, 691–702. [CrossRef] [PubMed]

37. Manahova, M.E.; Mostert, P.; Kok, P.; Schoffelen, J.M.; de Lange, F.P. Stimulus Familiarity and Expectation Jointly Modulate Neural Activity in the Visual Ventral Stream. *J. Cogn. Neurosci.* **2018**, *30*, 1366–1377. [CrossRef] [PubMed]
38. Rungratsameetaweemana, N.; Itthipuripat, S.; Salazar, A.; Serences, J.T. Expectations Do Not Alter Early Sensory Processing during Perceptual Decision-Making. *J. Neurosci.* **2018**, *38*, 5632–5648. [CrossRef] [PubMed]
39. Solomon, S.S.; Tang, H.; Sussman, E.; Kohn, A. Limited Evidence for Sensory Prediction Error Responses in Visual Cortex of Macaques and Humans. *Cereb. Cortex* **2021**, *31*, 3136–3152. [CrossRef]
40. Hall, M.G.; Mattingley, J.B.; Dux, P.E. Electrophysiological correlates of incidentally learned expectations in human vision. *J. Neurophysiol.* **2018**, *119*, 1461–1470. [CrossRef] [PubMed]
41. den Ouden, C.; Zhou, A.; Mepani, V.; Kovacs, G.; Vogels, R.; Feuerriegel, D. Stimulus expectations do not modulate visual event-related potentials in probabilistic cueing designs. *Neuroimage* **2023**, *280*, 120347. [CrossRef] [PubMed]
42. den Ouden, H.E.; Daunizeau, J.; Roiser, J.; Friston, K.J.; Stephan, K.E. Striatal prediction error modulates cortical coupling. *J. Neurosci.* **2010**, *30*, 3210–3219. [CrossRef] [PubMed]
43. Egner, T.; Monti, J.M.; Summerfield, C. Expectation and surprise determine neural population responses in the ventral visual stream. *J. Neurosci.* **2010**, *30*, 16601–16608. [CrossRef]
44. Richter, D.; Ekman, M.; de Lange, F.P. Suppressed Sensory Response to Predictable Object Stimuli throughout the Ventral Visual Stream. *J. Neurosci.* **2018**, *38*, 7452–7461. [CrossRef] [PubMed]
45. Richter, D.; de Lange, F.P. Statistical learning attenuates visual activity only for attended stimuli. *eLife* **2019**, *8*, e47869. [CrossRef] [PubMed]
46. Summerfield, C.; de Lange, F.P. Expectation in perceptual decision making: Neural and computational mechanisms. *Nat. Rev. Neurosci.* **2014**, *15*, 745–756. [CrossRef] [PubMed]
47. Duncan, D.; Theeuwes, J. Statistical learning in the absence of explicit top-down attention. *Cortex* **2020**, *131*, 54–65. [CrossRef] [PubMed]
48. St John-Saaltink, E.; Utzerath, C.; Kok, P.; Lau, H.C.; de Lange, F.P. Expectation Suppression in Early Visual Cortex Depends on Task Set. *PLoS ONE* **2015**, *10*, e0131172. [CrossRef] [PubMed]
49. Auksztulewicz, R.; Schwiedrzik, C.M.; Thesen, T.; Doyle, W.; Devinsky, O.; Nobre, A.C.; Schroeder, C.E.; Friston, K.J.; Melloni, L. Not All Predictions Are Equal: "What" and "When" Predictions Modulate Activity in Auditory Cortex through Different Mechanisms. *J. Neurosci.* **2018**, *38*, 8680–8693. [CrossRef]
50. Moskowitz, H.S.; Sussman, E.S. Sound category habituation requires task-relevant attention. *Front. Neurosci.* **2023**, *17*, 1228506. [CrossRef] [PubMed]
51. Stokes, M.G.; Myers, N.E.; Turnbull, J.; Nobre, A.C. Preferential encoding of behaviorally relevant predictions revealed by EEG. *Front. Hum. Neurosci.* **2014**, *8*, 687. [CrossRef] [PubMed]
52. Brogden, W.J. Sensory pre-conditioning of human subjects. *J. Exp. Psychol.* **1947**, *37*, 527–539. [CrossRef] [PubMed]
53. Chernikoff, R.; Brogden, W.J. The effect of instructions upon sensory preconditioning of human subjects. *J. Exp. Psychol.* **1949**, *39*, 200–207. [CrossRef] [PubMed]
54. Headley, D.B.; Weinberger, N.M. Relational associative learning induces cross-modal plasticity in early visual cortex. *Cereb. Cortex* **2015**, *25*, 1306–1318. [CrossRef]
55. Hoffeld, D.R.; Kendall, S.B.; Thompson, R.F.; Brogden, W.J. Effect of amount of preconditioning training upon the magnitude of sensory preconditioning. *J. Exp. Psychol.* **1960**, *59*, 198–204. [CrossRef] [PubMed]
56. Etzel, J.A.; Courtney, Y.; Carey, C.E.; Gehred, M.Z.; Agrawal, A.; Braver, T.S. Pattern Similarity Analyses of FrontoParietal Task Coding: Individual Variation and Genetic Influences. *Cereb. Cortex* **2020**, *30*, 3167–3183. [CrossRef] [PubMed]
57. Sommer, V.R.; Mount, L.; Weigelt, S.; Werkle-Bergner, M.; Sander, M.C. Spectral pattern similarity analysis: Tutorial and application in developmental cognitive neuroscience. *Dev. Cogn. Neurosci.* **2022**, *54*, 101071. [CrossRef]
58. Walther, A.; Nili, H.; Ejaz, N.; Alink, A.; Kriegeskorte, N.; Diedrichsen, J. Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage* **2016**, *137*, 188–200. [CrossRef]
59. Lieder, F.; Daunizeau, J.; Garrido, M.I.; Friston, K.J.; Stephan, K.E. Modelling trial-by-trial changes in the mismatch negativity. *PLoS Comput. Biol.* **2013**, *9*, e1002911. [CrossRef] [PubMed]
60. Lieder, F.; Stephan, K.E.; Daunizeau, J.; Garrido, M.I.; Friston, K.J. A neurocomputational model of the mismatch negativity. *PLoS Comput. Biol.* **2013**, *9*, e1003288. [CrossRef]
61. Mathys, C.D.; Lomakina, E.I.; Daunizeau, J.; Iglesias, S.; Brodersen, K.H.; Friston, K.J.; Stephan, K.E. Uncertainty in perception and the Hierarchical Gaussian Filter. *Front. Hum. Neurosci.* **2014**, *8*, 825. [CrossRef] [PubMed]
62. Mathot, S.; Schreij, D.; Theeuwes, J. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* **2012**, *44*, 314–324. [CrossRef] [PubMed]
63. Delorme, A.; Makeig, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **2004**, *134*, 9–21. [CrossRef] [PubMed]
64. Gabard-Durnam, L.J.; Mendez Leal, A.S.; Wilkinson, C.L.; Levin, A.R. The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): Standardized Processing Software for Developmental and High-Artifact Data. *Front. Neurosci.* **2018**, *12*, 97. [CrossRef]

65. Lopez, K.L.; Monachino, A.D.; Morales, S.; Leach, S.C.; Bowers, M.E.; Gabard-Durnam, L.J. HAPPILEE: HAPPE In Low Electrode Electroencephalography, a standardized pre-processing software for lower density recordings. *Neuroimage* **2022**, *260*, 119390. [CrossRef] [PubMed]
66. Bell, A.J.; Sejnowski, T.J. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **1995**, *7*, 1129–1159. [CrossRef] [PubMed]
67. Pion-Tonachini, L.; Kreutz-Delgado, K.; Makeig, S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* **2019**, *198*, 181–197. [CrossRef]
68. Oostenveld, R.; Fries, P.; Maris, E.; Schoffelen, J.M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, *2011*, 156869. [CrossRef] [PubMed]
69. Myers, N.E.; Stokes, M.G.; Walther, L.; Nobre, A.C. Oscillatory brain state predicts variability in working memory. *J. Neurosci.* **2014**, *34*, 7735–7743. [CrossRef] [PubMed]
70. Treder, M.S. MVPA-Light: A Classification and Regression Toolbox for Multi-Dimensional Data. *Front. Neurosci.* **2020**, *14*, 289. [CrossRef] [PubMed]
71. Maris, E.; Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **2007**, *164*, 177–190. [CrossRef] [PubMed]
72. Pernet, C.R.; Latinus, M.; Nichols, T.E.; Rousselet, G.A. Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *J. Neurosci. Methods* **2015**, *250*, 85–93. [CrossRef] [PubMed]
73. Haufe, S.; Meinecke, F.; Gorgen, K.; Dahne, S.; Haynes, J.D.; Blankertz, B.; Biessmann, F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* **2014**, *87*, 96–110. [CrossRef] [PubMed]
74. Mathys, C.; Daunizeau, J.; Friston, K.J.; Stephan, K.E. A bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* **2011**, *5*, 39. [CrossRef] [PubMed]
75. Hauser, T.U.; Iannaccone, R.; Ball, J.; Mathys, C.; Brandeis, D.; Walitza, S.; Brem, S. Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry* **2014**, *71*, 1165–1173. [CrossRef] [PubMed]
76. Stefanics, G.; Heinzle, J.; Horvath, A.A.; Stephan, K.E. Visual Mismatch and Predictive Coding: A Computational Single-Trial ERP Study. *J. Neurosci.* **2018**, *38*, 4020–4030. [CrossRef] [PubMed]
77. Rescorla, R.A.; Wagner, A. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non reinforcement. In *Classical Conditioning II: Current Research and Theory*; Black, A.H., Prokasy, W.F., Eds.; Appleton-Century-Crofts: New York, NY, USA, 1972; pp. 64–99.
78. Bueti, D.; Macaluso, E. Auditory temporal expectations modulate activity in visual cortex. *Neuroimage* **2010**, *51*, 1168–1183. [CrossRef] [PubMed]
79. McDonald, J.J.; Stormer, V.S.; Martinez, A.; Feng, W.; Hillyard, S.A. Salient sounds activate human visual cortex automatically. *J. Neurosci.* **2013**, *33*, 9194–9201. [CrossRef] [PubMed]
80. Garner, A.R.; Keller, G.B. A cortical circuit for audio-visual predictions. *Nat. Neurosci.* **2022**, *25*, 98–105. [CrossRef]
81. Fiser, J.; Lengyel, G. A common probabilistic framework for perceptual and statistical learning. *Curr. Opin. Neurobiol.* **2019**, *58*, 218–228. [CrossRef] [PubMed]
82. Watanabe, T.; Sasaki, Y. Perceptual learning: Toward a comprehensive theory. *Annu. Rev. Psychol.* **2015**, *66*, 197–221. [CrossRef] [PubMed]
83. Rahnev, D.; Lau, H.; de Lange, F.P. Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. *J. Neurosci.* **2011**, *31*, 10741–10748. [CrossRef] [PubMed]
84. Giustino, T.F.; Maren, S. The Role of the Medial Prefrontal Cortex in the Conditioning and Extinction of Fear. *Front. Behav. Neurosci.* **2015**, *9*, 298. [CrossRef] [PubMed]
85. Kim, J.; Ghim, J.W.; Lee, J.H.; Jung, M.W. Neural correlates of interval timing in rodent prefrontal cortex. *J. Neurosci.* **2013**, *33*, 13834–13847. [CrossRef] [PubMed]
86. Sznabel, D.; Land, R.; Kopp, B.; Kral, A. The relation between implicit statistical learning and proactivity as revealed by EEG. *Sci. Rep.* **2023**, *13*, 15787. [CrossRef] [PubMed]
87. Donchin, E.; Heffley, E.; Hillyard, S.A.; Loveless, N.; Maltzman, I.; Ohman, A.; Rosler, F.; Ruchkin, D.; Siddle, D. Cognition and event-related potentials. II. The orienting reflex and P300. *Ann. N. Y. Acad. Sci.* **1984**, *425*, 39–57. [CrossRef] [PubMed]
88. Stefanics, G.; Kremlacek, J.; Czigler, I. Visual mismatch negativity: A predictive coding view. *Front. Hum. Neurosci.* **2014**, *8*, 666. [CrossRef]
89. Arnal, L.H.; Wyart, V.; Giraud, A.L. Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci.* **2011**, *14*, 797–801. [CrossRef] [PubMed]
90. Barraclough, N.E.; Xiao, D.; Baker, C.I.; Oram, M.W.; Perrett, D.I. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J. Cogn. Neurosci.* **2005**, *17*, 377–391. [CrossRef] [PubMed]
91. Friston, K. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2005**, *360*, 815–836. [CrossRef] [PubMed]
92. Donchin, E.; Coles, M.G.H. Is the P300 Component a Manifestation of Context Updating. *Behav. Brain Sci.* **1988**, *11*, 357–374. [CrossRef]
93. Verleger, R. Event-Related Potentials and Memory—A Critique of the Context Updating Hypothesis and an Alternative Interpretation of P3. *Behav. Brain Sci.* **1988**, *11*, 343–3568. [CrossRef]

94. Schwartenbeck, P.; Passecker, J.; Hauser, T.U.; FitzGerald, T.H.; Kronbichler, M.; Friston, K.J. Computational mechanisms of curiosity and goal-directed exploration. *eLife* **2019**, *8*, e41703. [CrossRef] [PubMed]
95. Schulz, E.; Gershman, S.J. The algorithmic architecture of exploration in the human brain. *Curr. Opin. Neurobiol.* **2019**, *55*, 7–14. [CrossRef] [PubMed]
96. Hebb, D.O. *The Organization of Behavior: A Neuropsychological Theory*; Wiley: New York, NY, USA, 1949; p. xix. 335p.
97. Rescorla, R.A. Pavlovian Conditioning—Its Not what You Think It Is. *Am. Psychol.* **1988**, *43*, 151–160. [CrossRef] [PubMed]
98. Rescorla, R.A. Behavioral-Studies of Pavlovian Conditioning. *Annu. Rev. Neurosci.* **1988**, *11*, 329–352. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.