



Differential effects of prediction error and adaptation along the auditory cortical hierarchy during deviance processing

Insa Schlossmacher^{a,b,*}, Jacky Dilly^a, Ina Protmann^a, David Hofmann^a, Torge Dellert^{a,b}, Marie-Luise Roth-Paysen^a, Robert Moeck^a, Maximilian Bruchmann^{a,b}, Thomas Straube^{a,b}

^a Institute of Medical Psychology and Systems Neuroscience, University of Münster, Münster 48149, Germany

^b Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Münster, Münster 48149, Germany

ARTICLE INFO

Keywords:

Adaptation
Prediction error
fMRI
Oddball
Deviance processing

ABSTRACT

Neural mismatch responses have been proposed to rely on different mechanisms, including prediction error-related activity and adaptation to frequent stimuli. However, the hierarchical cortical structure of these mechanisms is unknown. To investigate this question, we recorded hemodynamic responses while participants (N = 54) listened to an auditory oddball sequence as well as a suited control condition. In addition to effects in sensory processing areas (Heschl's gyrus, superior temporal gyrus (STG)), we found several distinct clusters that indexed deviance processing in frontal and parietal regions (anterior cingulate cortex/supplementary motor area (ACC/SMA), inferior parietal lobule (IPL), anterior insula (AI), inferior frontal junction (IFJ)). Comparing responses to the control stimulus with the deviant and standard enabled us to delineate the contributions of prediction error- or adaptation-related brain activation, respectively. We observed significant effects of adaptation in Heschl's gyrus, STG and ACC/SMA, while prediction error-related activity was observed in STG, IPL, AI and IFJ. Additional dynamic causal modeling confirmed the superiority of a hierarchical processing structure compared to a flat structure. Thus, we found that while prediction-error related processes increased with the hierarchical level of the brain area, adaptation declined. This suggests that the relative contribution of different mechanisms in deviance processing varies across the cortical hierarchy.

1. Introduction

Detecting changes in our environment is a cornerstone of perception. A vast amount of research effort has been put into the investigation of neural correlates of deviance processing including neuroimaging (Kim, 2014) and electrophysiological approaches (Näätänen et al., 2011; Polich, 2007; Stefanics et al., 2014), showing increased activity to deviant stimuli in different variants of oddball designs. Recently, hierarchical predictive processing has been put forward as a theoretical underpinning for these deviance-related effects (Clark, 2013; Garrido et al., 2009b; Stefanics et al., 2014; Winkler and Czigler, 2012). From this point of view, the increase of neural activation for unexpected rare stimuli compared to expected frequent ones stems from a process where a prediction is compared with the actual sensory input. If prediction and input do not match, as is the case when a rare – and thus unexpected – deviant stimulus is presented, a prediction error signal would be elicited. This prediction error signal would then be propagated upwards in the hierarchy and compared with the predictions of the next higher level and so forth, enabling efficient information processing (Clark, 2013; Stefanics et al., 2014).

However, while predictive processing offers a compelling explanation for deviance-related effects by means of prediction errors, it is not the only process that could be responsible for an observed difference between rare and frequent stimuli. Stimulus-specific adaptation (SSA) has also been proposed to play an important role during deviance processing (Jääskeläinen et al., 2004; May and Tiitinen, 2010; Nelken, 2014). In this theoretical framework, the difference between rare and frequent stimuli stems from habituation of neuronal responsiveness to the frequent stimulus, which thus elicits a smaller response compared to the non-adapted cells (fresh afferents) activated by the rare stimulus. In other words, deviance responses in typical oddball designs are not driven by genuine mismatch responses but by altered, i.e., reduced responses to the standard stimulus. While most work on SSA relies on electrophysiological studies in animals, in humans, repetition suppression (RS) or fMRI adaptation (Barron et al., 2016; Larsson et al., 2016) represents a similar paradigm where neural activity after a stimulus repetition is reduced (Aukstulewicz and Friston, 2016; Barron et al., 2016). Similar to deviance responses in oddball paradigms, mechanisms relying on neuronal fatigue and predictive coding have been suggested

* Corresponding author at: Institute of Medical Psychology and Systems Neuroscience, University of Münster, Von-Esmarch-Strasse 52, Münster 48149, Germany.
E-mail address: insa.schlossmacher@uni-muenster.de (I. Schlossmacher).

to underlie repetition suppression (Auzstulewicz and Friston, 2016; Barron et al., 2016).

At first glance, the prediction error and adaptation account seem opposing. However, there is evidence that both processes can be present in the brain to variable degrees in the same region or at the same time (Ishishita et al., 2019; Laufer et al., 2008; Opitz et al., 2005; Parras et al., 2017). It is also important to note that both adaptation and prediction can be reconciled in the predictive processing framework (Garrido et al., 2009b, 2009a; O'Shea, 2015). From this point of view, adaptation could be considered a lower-level predictive process that contributes through synaptic changes to the precision, while higher-level processes evidenced by prediction errors define the flow of information between cortical areas (Garrido et al., 2009b, 2009a). Thus, from this unified perspective, observing adaptation or prediction as a mechanism in MMN generation gives important insights on the level and complexity of processing. One way to delineate the contribution of these two mechanisms to deviance effects is to compare the responses to deviant and standard stimuli to a control stimulus (Maess et al., 2007; Parras et al., 2017; Schröger and Wolff, 1996). While standard and deviant stimuli usually differ on a physical dimension, e.g., frequency (Näätänen et al., 2011, 2007), the control stimulus is in most cases physically identical to the deviant and presented in a multi-standard paradigm where all stimuli have the same stimulus probability as the deviant. Comparing the deviant with the control thus offers a way to distill prediction error activity without being contaminated by adaptation. Additionally, comparing the control with the standard stimulus shows adaptation-related activity.

Using this experimental procedure, there is initial evidence from an electrophysiological study in rodents (Parras et al., 2017) that the relative contribution of prediction error vs. adaptation increases from subcortical areas to primary auditory cortex as well as from primary auditory cortex to the posterior auditory field (Parras et al., 2021). In line with these results, two neuroimaging studies specifically targeting the auditory cortex showed the presence of both mechanisms within Heschl's gyrus and superior temporal gyrus (STG) with an anterior-posterior gradient in humans (Laufer et al., 2008; Opitz et al., 2005). Furthermore, an electrocorticography (ECoG) study found that temporal areas were related to predictable changes, while frontal areas indexed unpredictable changes (Dürschmid et al., 2016). However, to date, there are no brain imaging studies that investigated the effects of prediction error vs. adaptation across all typical cortical brain areas commonly showing deviance-related effects (Kiehl et al., 2005; Kim, 2014). Besides primary and secondary auditory cortex, encompassing Heschl's gyrus and STG, several frontal and parietal areas included in ventral and dorsal attention networks, as well as subcortical areas, reliably show increased activation to deviant vs. standard stimuli (Kiehl et al., 2005; Kim, 2014). Thus, deviance responses are processed on different hierarchical levels traversed during auditory processing from the thalamus to Heschl's gyrus and STG and onwards to higher order association cortices (Li et al., 2019; Parras et al., 2017). In the context of the current study, we consider subcortical regions and regions linked with sensory processing as possessing a lower hierarchical level and cortical regions not directly linked with sensory processing a higher level (Garrido et al., 2009a, 2008; Li et al., 2019; Parras et al., 2017).

While the investigation of mechanisms underlying deviance responses has been researched intensively using electrophysiological methods (see, e.g., Ruhnau et al., 2012; Wiens et al., 2019), the hierarchical dimension of deviance processing, i.e., whether and how mechanisms vary depending on the cortical region, has not been investigated yet. In order to address this gap, we conducted a high-powered fMRI study ($N = 54$) using an oddball design with a suited control condition that allowed to delineate prediction error- and adaptation-related mechanisms in typical areas involved in auditory deviance processing in humans. Following prior research (Dürschmid et al., 2016; Parras et al., 2017), we hypothesized that on lower levels of the cortical hierarchy, adaptation-related activity will be more prominent, while on higher levels, activity should be driven by prediction errors.

2. Methods

2.1. Participants

Fifty-nine right-handed participants with normal hearing and no history of neurological or psychiatric illness took part in the experiment and were compensated with €10/h. Four participants had to be excluded due to excessive head movements (> 3 mm) during recording and one participant because of anatomical MRI anomalies. The remaining 54 participants (39 female) were aged from 18 to 33 ($M = 23.20$, $SD = 3.04$). The local ethics committee has approved the study and all procedures were carried out in accordance with the Helsinki declaration.

2.2. Experimental procedure and stimulus material

Stimuli consisted of pure sine-tones of 600, 800, 1000, 1200, and 1400 Hz. Stimulus duration amounted to 100 ms, including rise/fall times of 10 ms. The sound volume was chosen to be easily audible during functional sequences but not unpleasantly loud. In oddball blocks, the 800 and 1200 Hz tones served as deviant and standard counterbalanced across participants. The probability of deviant to standard stimuli was 20:80 and stimulus presentation was pseudorandomized so that no two deviants were presented consecutively. In control blocks, all five tone stimuli were presented randomly with a probability of 20% (see Fig. 1A). Throughout, participants' task was to respond to target stimuli of 300 ms duration, which were randomly interspersed in the sequence. In total, 40 targets were included, the frequency of which conformed to the stimulus probability of the current block, i.e., all tone frequencies could be potential targets. We included this task in order to ensure that participants attended to the tone sequence but did not view the deviant as a target. Interstimulus intervals (ISIs) ranged from 1.04 to 17.25 s ($M = 3.01$, $SD = 2.06$) and were derived using optseq2 (Greve, 2009). Four different optseq sequences were computed and randomly assigned to two runs per participant, creating 12 different sequence combinations. In total, two runs of 250 stimuli (ca. 13 min each) were acquired and separated by a short break. One run could start with either an oddball or control block, which seamlessly changed to control or oddball block in the middle of the run. Whether the participants started with the oddball or control block was counterbalanced across participants. The second run was always presented in reverse order, e.g., if the first run was oddball-control, the second run was control-oddball. Before starting the experiment, a short practice block of about 1 min was presented in order to accustom the participants to the task. Stimuli in this practice block were structured like the experimental block that followed, i.e., they also included brief oddball and control sequences. At all times, a white fixation cross was presented on a black screen, and participants were asked to fixate during the run. Stimulus presentation and response collection were controlled by the software Presentation (version 21.1, Neurobehavioral Systems, Albany, CA).

2.3. Data acquisition and preprocessing

A 3-Tesla Siemens Magnetom Prisma with a 20-channel Siemens Head Matrix Coil (Siemens Medical Systems, Erlangen, Germany) was used to acquire MRI data. In a first step, we obtained a high-resolution T1-weighted scan with 192 slices for anatomical localization and coregistration (repetition time (TR) = 2130 ms, echo time (TE) = 2.28 ms, flip angle (FA) = 8°, field of view (FOV) = 256 × 256 mm, voxel size = 1 × 1 × 1 mm). A shimming field was applied in order to minimize magnetic field inhomogeneity. Then, we recorded two functional datasets per participant (2 runs) consisting of 353 volumes and 42 slices each by means of a T2*-weighted echoplanar sequence sensitive to blood oxygenation level-dependent (BOLD) contrast (TR = 2300 ms, TE = 30 ms, FA = 90°, FOV = 216 × 216 mm, voxel size = 3 × 3 × 3 mm).

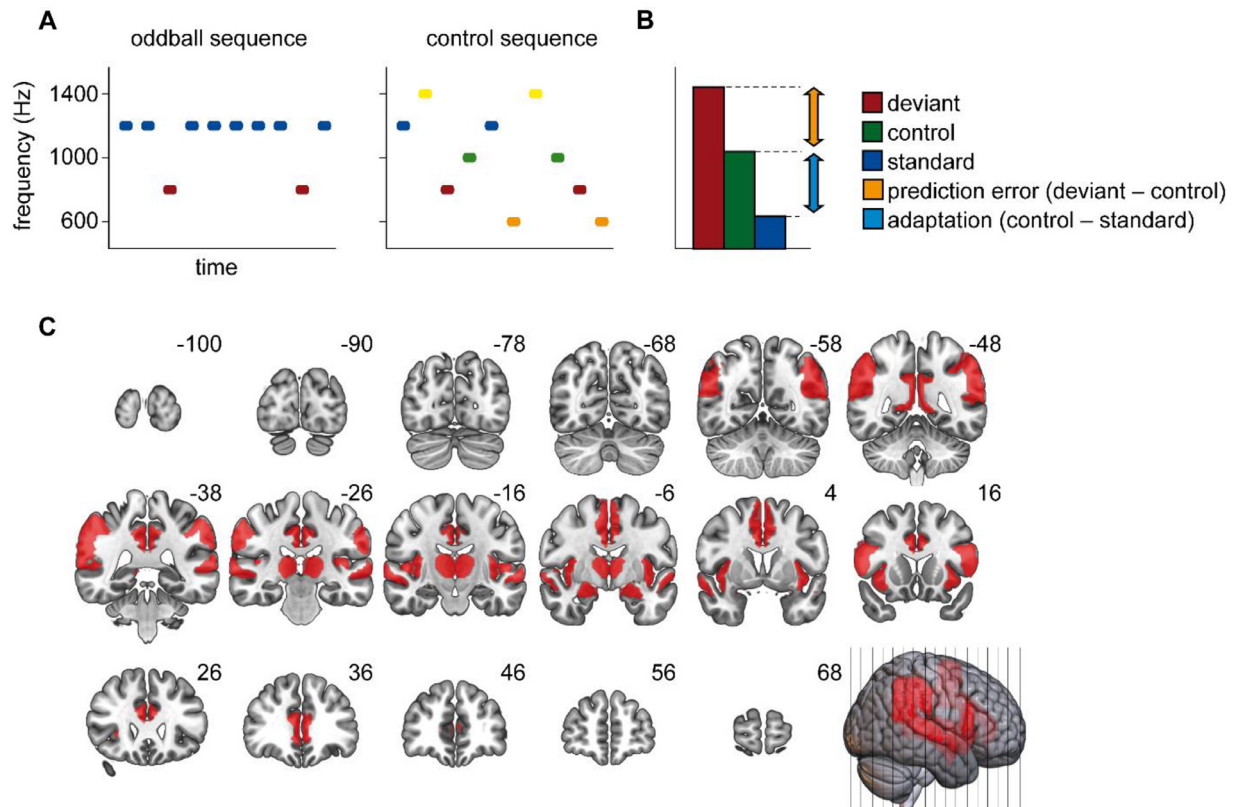


Fig. 1. Experimental paradigm and mask for data analysis. (A) Schematic of oddball and control sequence. In the experiment, the frequencies of deviant and standard stimulus were counterbalanced across participants. (B) Decomposition of observed responses into prediction error-related and adaptation-related activity. (C) Illustration of the mask applied during the cluster-based permutation. Red areas were included in the mask.

Preprocessing relied on SPM12 v7771 (Wellcome Department of Cognitive Neurology, London, UK) and the Data Processing & Analysis of Brain Imaging (DPABI) 4.3 toolbox (Yan et al., 2016) in MATLAB. We removed the first five data volumes to account for spin saturation effects. Then, slice-scan-time correction and realignment using a six-parameter (rigid body) linear transformation was performed. In a next step, we co-registered anatomical and functional images and segmented these into gray matter, white matter and cerebrospinal fluid. Finally, we normalized functional data to Montreal Neurological Institute (MNI) standard space using DARTEL (Ashburner, 2007), resampled it to 3 mm isotropic voxels and spatially smoothed it with an 8 mm full width at half maximum Gaussian kernel.

2.4. Statistical analysis

A general linear model (GLM) was estimated for each participant in the first-level analysis. In order to eliminate slow signal drifts, we used a high-pass filter with a cutoff of 128 seconds. We applied SPM's pre-whitening method FAST (Corbin et al., 2018) to model autocorrelations as recommended by Olszowy et al. (2019). The GLM design matrix contained the onsets of deviants, standards, and controls. Onsets of controls were divided up into two separate predictors. One used the onsets of the stimulus physically identical to the deviant (later compared with the deviant and standard); the other modeled the onsets of all other control stimuli. The predictor for the remaining controls as well as predictors for targets, responses and six head movement parameters were included in the first-level GLM as nuisance regressors. Onsets were convolved with a 2-gamma hemodynamic response function to model the BOLD signal change for each predictor. Contrast images (deviant – standard) of the beta estimates were created for each participant for the second-level analysis.

In order to isolate mismatch-related activity in the second-level analysis, we used cluster-based permutation (CBP) as implemented in PALM (Winkler et al., 2014). In this approach, t -tests are computed in each voxel, and then all neighboring t -values that cross a predefined threshold (voxel-wise α) are summed up. This sum is referred to as the mass of a cluster, which is then compared with the distribution of cluster masses under the null hypothesis derived from random permutations of the data. CBPs offer a nonparametric statistical test while circumventing the multiple comparison problem (Eklund et al., 2016). The voxel-wise α amounted to .001, and a cluster was deemed significant with $\alpha < .05$. The number of permutations was set to 10000.

In order to maximize the statistical power of the CBP, we opted for a region-of-interest (ROI) analysis in which we included areas that indexed deviance processing in prior studies based on the meta-analytic results of Kim (2014). The following areas of interest were identified and included in one mask based on the Harvard Oxford Atlas (Desikan et al., 2006): Heschl's gyrus, superior temporal gyrus (STG), anterior and posterior cingulate cortex (ACC/PCC), supplementary motor area (SMA), inferior frontal junction (IFJ), inferior parietal lobule (IPL), temporo-parietal junction (TPJ), insula, thalamus and amygdala (see Fig. 1C). Areas were chosen to correspond to the modality and task of the current study (auditory and task-irrelevant, see Kim, 2014 for details). Given a significant cluster, averaged betas were extracted from all voxels within the group-level cluster for deviant, standard and control stimuli, averaged across hemispheres, and z-standardized separately for each region to account for differences in raw beta values. For clusters comprising several distinct regions, in our case, Heschl's gyrus and STG, we extracted the cluster averages using only the significant voxels that were encompassed in the original HOA mask templates. That is, we extracted the betas for Heschl's gyrus and STG separately. Then, adaptation-related activity was computed as control – standard and prediction error-related activity as deviant – control (see Fig. 1B). These differ-

ences were then subjected to a 2×6 repeated-measures ANOVA with the factors mechanism (adaptation, prediction error) and region (Heschl's gyrus, STG, ACC/SMA, IPL, AI, IFJ) to check whether adaptation-related and prediction error-related activity varied with brain region. Both factors vary within subjects and thus represent repeated measures. Violations of sphericity were corrected using the Greenhouse-Geisser procedure and corrected p -values as well as $\hat{\epsilon}$ -values are reported. Two-sided t -tests were used to follow up the significant interactions of mechanism with brain region. We used Bonferroni correction to correct for the multiple tests. The α level was set to 0.008 for testing prediction and adaptation (6 tests each), respectively, and to 0.01 for comparing the difference of the mechanisms between one region and the others (5 tests each). We additionally report Bayes Factors (BF), with BF_{01} denoting the evidence for the null hypothesis (either prediction error = deviant – control = 0 or adaptation = control – standard = 0) and BF_{10} the evidence for the alternative hypothesis (either prediction error = deviant – control \neq 0 or adaptation = control – standard \neq 0). We use the conventions from [Jeffreys \(1961\)](#) to interpret the results of our Bayesian analyses.

In order to investigate whether the six brain areas found in the main analysis were organized in a hierarchical or flat structure, we used dynamic causal modeling (DCM; [Friston et al., 2003](#)). The volume-of-interest tool included in SPM 12 was used to extract ROI time series from the unsmoothed concatenated data. To this end, we built masks including 8 mm spheres around bilateral group-level peaks. Following recommendations by [Zeidman et al. \(2019a\)](#), we used the first principal component and controlled for the F -contrast using a liberal voxelwise α of 0.1 based on the main effect of oddball vs. standard. Two participants failed to show activity in all six regions and were excluded from the analysis.

For the remaining 52 participants, we defined two DCMs representing a flat and a hierarchical structure which only differed in the way the oddball modulated the connections (see [Fig. 4A](#)). We used a DCM for time series analysis ([Friston et al., 2003](#)). For both models, we modeled standard, oddball, and control as inputs to Heschl's gyrus based on physiological considerations. We defined a full model with all connections as baseline (matrix A) and varied how connections were modulated by the oddball (matrix B). For the hierarchical model, we assumed that Heschl's gyrus mainly connects to STG, and STG to higher-order regions like the IFJ (see, e.g., [Garrido et al., 2009a, 2008](#)). Based on the results of [Li et al. \(2019\)](#), the AI could be considered an intermediate node between STG and IFJ, while the ACC/SMA could be connected to the STG. While we did not have specific hypotheses concerning the IPL from previous research with oddball designs, we would assume that it also represents a higher level in the processing of deviant sounds maybe also being connected to the STG ([Binkofski et al., 2016](#)). The alternative model had a flat design, where Heschl's gyrus is connected to all other regions. All other configurations were left at default values and equal for both models. After estimating these DCMs at the participant level, Bayesian model comparison was performed on the second level using *spm_dcm_bmc_peb*, testing parameter estimates of matrix B ([Zeidman et al., 2019b](#)).

2.5. Data and code availability statement

Data and code are available on the Open Science Framework accessible via <https://osf.io/eufd6/>.

3. Results

3.1. Behavioral data

Performance on the duration task was very high, indexed by an average hit rate of 0.92 (SD = 0.15), an average false-alarm rate of 0.003 (SD = 0.007), and an average d' of 4.59 (SD = 0.77), indicating that participants were able to comply with the task easily. Furthermore, performance did not differ significantly between the oddball and control

sequence (hit rate: $t(53) = 0.66, p = .51, BF_{01} = 5.46$; false alarm rate: $t(53) = -0.81, p = .42, BF_{01} = 4.95$; d' : $t(53) = 0.15, p = .89, BF_{01} = 6.66$).

3.2. fMRI data

3.2.1. Main effects of mismatch processing

The cluster-based permutation revealed bilateral clusters of significant mismatch processing in the auditory cortex, including STG and Heschl's gyrus (right: $p < .001$; left: $p < .001$), the ACC/SMA (right: $p = .01$; left: $p = .002$), the IFJ (right: $p = .003$; left: $p = .002$) and the AI (right: $p = .008$; left: $p = .006$). Furthermore, a significant cluster was found in the left IPL ($p = .009$). Please see [Fig. 2](#) for a visualization of clusters and beta-values and [Table 1](#) for peak coordinates, t -statistics, and number of voxels (k) of the effects.

3.2.2. Mechanisms of mismatch processing

We found that the contribution of the different mechanisms, i.e., adaptation (control – standard) and prediction error (deviant – control), varied depending on the brain region. The repeated-measures ANOVA on the differential activity indicated a significant main effect of area ($F(5,265) = 2.54, p = .045, \hat{\epsilon} = 0.75$) and a significant interaction of area and mechanism ($F(5,265) = 3.28, p = .01, \hat{\epsilon} = 0.79$), while the main effect of mechanism did not reach significance ($F(1,53) = 0.03, p = .86$). We found significant adaptation in Heschl's gyrus ($t(53) = 4.64, p < .001, BF_{10} = 847.85$), STG ($t(53) = 4.52, p < .001, BF_{10} = 588.07$) and ACC/SMA ($t(53) = 2.84, p = .006, BF_{10} = 5.41$), while in the AI ($t(53) = 1.81, p = .08, BF_{01} = 1.49$), IPL ($t(53) = 1.08, p = .29, BF_{01} = 3.90$) and IFJ ($t(53) = 1.57, p = .12, BF_{01} = 2.14$), adaptation did not reach significance (see [Fig. 3A](#)). In contrast, we observed prediction error-related activity in the STG ($t(53) = 3.02, p = .004, BF_{10} = 8.36$), AI ($t(53) = 3.75, p < .001, BF_{10} = 57.97$), IPL ($t(53) = 2.93, p = .005, BF_{10} = 6.67$) and IFJ ($t(53) = 3.18, p = .002, BF_{10} = 12.42$), while no significant prediction error was found in Heschl's gyrus ($t(53) = 1.39, p = .17, BF_{01} = 2.71$) and ACC/SMA ($t(53) = 1.42, p = .16, BF_{01} = 2.62$).

In order to test how the relative contribution of the mechanisms differs between areas, we computed the difference between adaptation-related and prediction error-related activity. Mechanisms in Heschl's gyrus differed significantly from AI ($t(53) = -3.34, p = .002, BF_{10} = 18.75$), and IFJ ($t(53) = -2.57, p = .01, BF_{10} = 2.89$). STG differed significantly from AI ($t(53) = -2.60, p = .01, BF_{10} = 3.13$). Differences between Heschl's gyrus and IPL ($t(53) = -2.48, p = .02, BF_{10} = 2.39$) as well as STG and IFJ ($t(53) = -2.38, p = .02, BF_{10} = 1.95$) failed to reach corrected significance (all other $p > .05$). See [Fig. 3B](#) for a visualization of differences between cortical regions.

3.2.3. Dynamic causal modelling

Results of our DCM analysis can be found in [Fig. 4](#) including parameter estimates, model probabilities, and free energy of the models. The model probability for the second level was 1 for the hierarchical model compared to the flat model. The higher model evidence for the hierarchical model is further illustrated by its higher (more positive) free energy.

4. Discussion

In this study, we investigated the contribution of different mechanisms to explain increased deviance-related brain activation in humans. We found a hierarchical increase of the relation of predictive mechanisms vs. adaptation from primary auditory cortex across secondary auditory cortex to higher frontal and parietal regions.

Our first analytical step, in which we compared oddball and standard stimuli, confirmed deviance-related activation in most of our ROIs. This includes bilateral Heschl's gyrus and STG, which are commonly linked to auditory processing. Heschl's gyrus encompasses the primary auditory cortex ([Costa et al., 2011](#)), while the STG is involved in higher-order auditory processing ([Binder et al., 2000](#)). The effects in both ar-

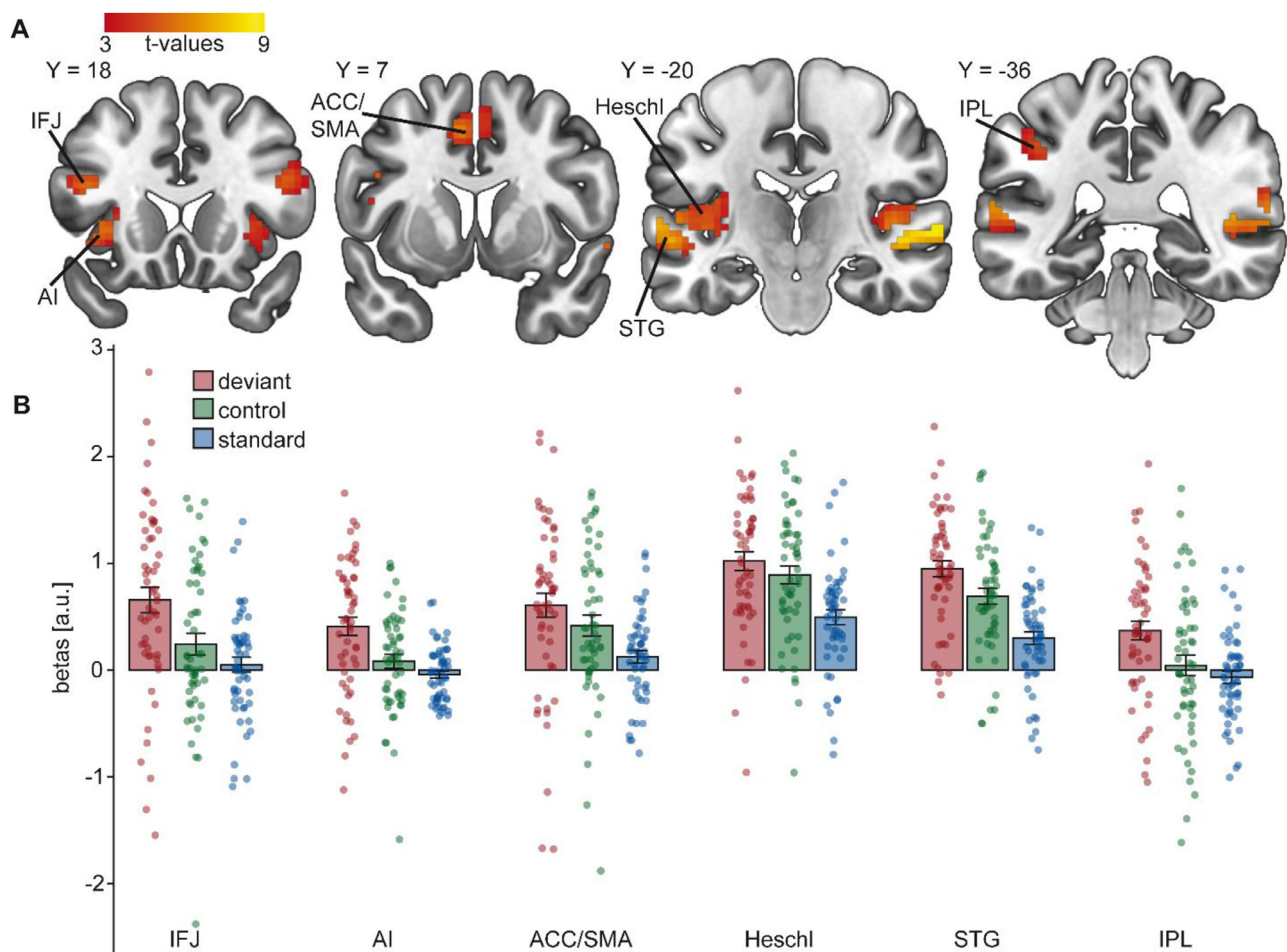


Fig. 2. Auditory mismatch responses in the brain. (A) Clusters found in the cluster-based permutation comparing deviant and standard stimulus. (B) Mean beta values extracted from the clusters found. The deviant, standard and control stimulus are displayed. Individual data points represent single subjects. ACC/SMA: anterior cingulate cortex/supplementary motor area, AI: anterior insula, IFJ: inferior frontal junction, IPL: inferior parietal lobule, STG: superior temporal gyrus.

Table 1

fMRI results of the oddball contrast. ACC/SMA: anterior cingulate cortex/supplementary motor area, AI: anterior insula, IFJ: inferior frontal junction, IPL: inferior parietal lobule, STG: superior temporal gyrus.

| Lobe | Area | Hemisphere | Peak MNI coordinates | | | t-statistics | | k |
|----------|---------|------------|----------------------|-----|----|--------------|---------|-----|
| | | | x | y | z | max(t) | mean(t) | |
| Temporal | STG | R | 63 | -18 | 0 | 9.51 | 5.63 | 286 |
| | | L | -54 | -3 | -6 | 7.80 | 5.34 | 264 |
| | Heschl | R | 51 | -9 | 0 | 5.87 | 4.54 | 90 |
| | | L | -57 | -15 | 6 | 6.03 | 4.68 | 131 |
| Frontal | IFJ | R | 45 | 15 | 21 | 5.38 | 4.25 | 74 |
| | | L | -42 | 9 | 24 | 6.94 | 4.58 | 83 |
| | ACC/SMA | R | 6 | 9 | 51 | 4.73 | 3.85 | 19 |
| | | L | -9 | 6 | 48 | 6.04 | 4.16 | 57 |
| Parietal | IPL | L | -39 | -39 | 39 | 5.10 | 4.08 | 41 |
| Insular | AI | R | 33 | 24 | 0 | 5.68 | 4.46 | 42 |
| | | L | -33 | 21 | 0 | 6.28 | 4.75 | 49 |

areas have been linked to the modulatory influences of the dorsal attention network (Kastner and Ungerleider, 2000), as well as preattentive change detection (Näätänen et al., 2007). Besides effects in sensory processing areas, we found several distinct clusters in frontal and parietal areas as well as the insula. Activations were mainly detected in the ventral attention network, which is involved in orienting attention to salient events and thus alerting the organism to environmental changes (Corbetta and Shulman, 2002; Sestieri et al., 2012). This includes the AI and ACC/SMA (Eckert et al., 2009; Yeo et al., 2011), which are also strongly involved in detecting salient events and initiating task-based at-

tention (Goulden et al., 2014; Menon and Uddin, 2010; Shackman et al., 2011). Furthermore, we found strong bilateral IFJ activity. The IFJ is involved in the dorsal fronto-parietal network modulating goal-directed attention in sensory areas in a top-down fashion (Kim, 2014; Yeo et al., 2011), but also in a network activated by unexpected salient environmental changes (Sestieri et al., 2012). This duality fits with the suggestion that the IFJ presents a dynamic region integrating information from both dorsal and ventral networks (Asplund et al., 2010). In addition to these activations, we also found deviance-related effects in the left IPL, which is part of a fronto-parietal control network (Cole et al., 2013;

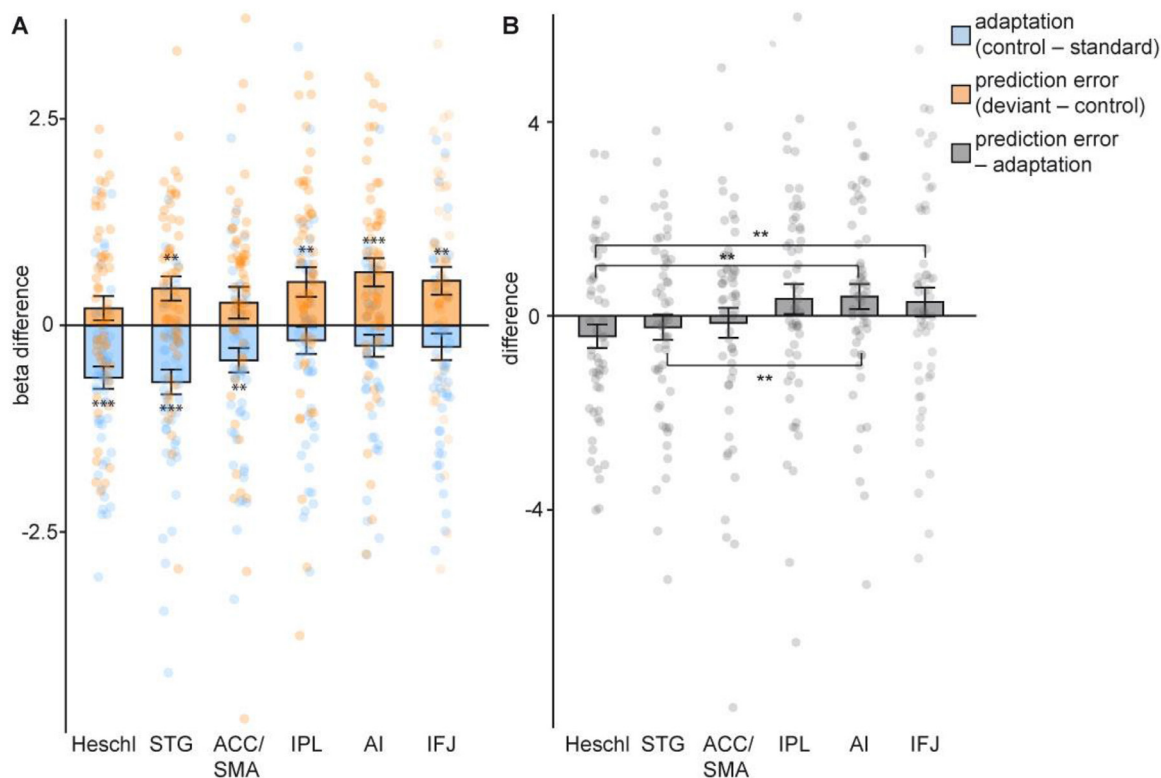


Fig. 3. Mechanisms of mismatch generation. (A) Adaptation and prediction in different brain regions. Prediction-error-related activity is computed by subtracting activity elicited by the control from the deviant. Adaptation-related activity is computed by subtracting activity elicited by the standard from the control. Note that positive prediction error-related activity is plotted upwards and positive adaptation-related activity is plotted downwards. (B) Relative contribution of the mechanisms, negative values correspond to a surplus of adaptation in the corresponding brain area, while positive values correspond to a surplus of prediction error. Error bars depict standard errors of the mean. Individual data points represent single subjects. IFJ: inferior frontal junction, AI: anterior insula, ACC/SMA: anterior cingulate cortex/supplementary motor area, STG: superior temporal gyrus, IPL: inferior parietal lobule. Asterisks correspond to the t -tests reported in the main text with * $p < .05$, ** $p < .01$, *** $p < .001$.

Corbetta and Shulman, 2002; Yeo et al., 2011), probably involved in indexing expectancy violations (O'Connor et al., 2010).

These deviance-related effects could be explained to varying degrees by adaptation- and prediction error-related activity. We found significant contributions of adaptation in Heschl's gyrus in line with previous research indicating sensory refractoriness effects in this area (Opitz et al., 2005), while contributions of prediction error did not reach significance. However, this does not mean that the primary auditory cortex does not show prediction error-related activation (Opitz et al., 2005), but that adaptation is – averaged across the deviance-related cluster – the primary driver of information processing. This point is further highlighted by the fact that we did not observe substantial Bayesian evidence for the absence of a prediction error effect (null hypothesis: prediction error equals zero). In contrast to primary auditory cortex, prediction error-related activity was significantly observed in the STG in accordance with previous studies showing an increase of predictive processes along the auditory processing hierarchy (Laufer et al., 2008; Opitz et al., 2005; Perras et al., 2021, 2017).

Furthermore, in the AI, IFJ and IPL, a significant prediction error effect was found, but no significant adaptation effect. However, BF's in favor of the null hypothesis (adaptation equals zero) indicated substantial evidence only in the IPL, while in the AI and the IFJ, anecdotal evidence for the null was observed. This suggests that while prediction error is the primary driver of activity, adaptation effects cannot be ruled out with confidence. Nonetheless, the finding of prediction error as the primary mechanism fits well with studies showing a dominant role of these regions during various predictive processes (Allen et al., 2016; Dürschmid et al., 2016; Geuter et al., 2017; O'Connor et al.,

2010; Siman-Tov et al., 2019). Surprisingly, we found no significant prediction error-related activity in ACC/SMA, probably due to the task-irrelevant oddball paradigm chosen here, which considerably differs from the paradigms linking ACC to predictive processing (Alexander and Brown, 2019). Furthermore, even though we used a large sample of participants, a further increase of sample size could probably alter results for ACC/SMA.

Comparing the relative contributions of prediction error and adaptation, we observed differences between areas mainly concerned with auditory processing and higher-order areas like the AI and IFJ. Thus, while prediction error is generated in auditory areas like the STG, its contribution to deviance-related effects increases on higher levels. These results are in line with the results of Dürschmid and colleagues (2016), who found an increase of predictive processing from temporal to frontal areas. Furthermore, this supposed hierarchical organization fits well with the results of Li and colleagues (2019). These authors reconstructed the temporal evolution of deviance-related responses by combining EEG and fMRI and found the information to flow from auditory cortex via the insula to the inferior frontal cortex. Our findings are also in line with the proposal of hierarchical prediction error processing as a basic principle in the human brain (Clark, 2013; Heilbron and Chait, 2018). In this framework, auditory cortices are said to form a generative model of the acoustical environment, which is informed by predictions from higher level areas (Heilbron and Chait, 2018). We did not observe equal amounts of prediction error in all regions involved in deviance processing. At first glance, this seems to disagree with predictive coding theories that hypothesize the occurrence of error and prediction units on all cortical levels (Clark, 2013; Heilbron and Chait, 2018). However, if

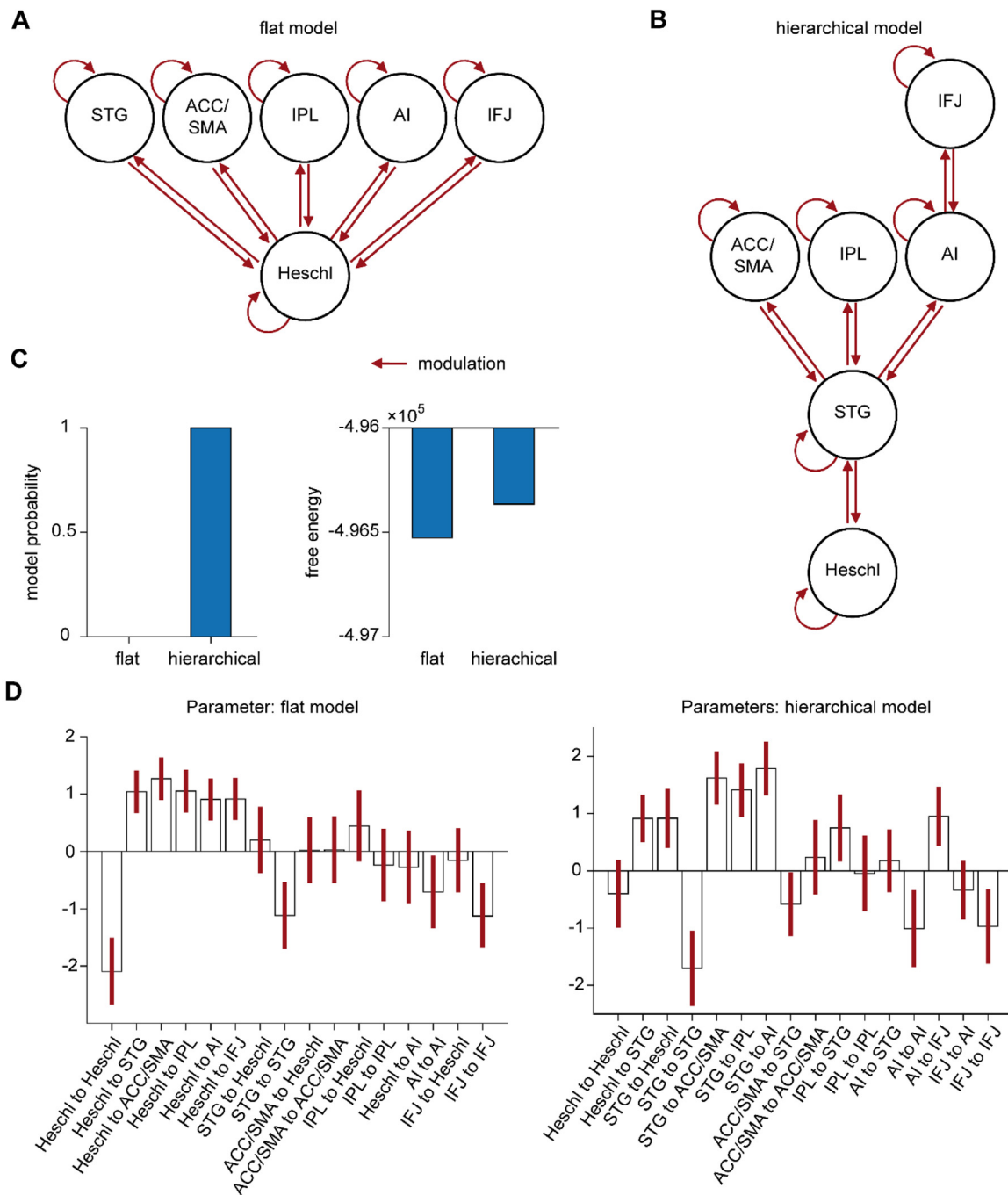


Fig. 4. DCM analysis. (A) Model architecture of the flat model. (B) Model architecture of the hierarchical model. (C) Results of Bayesian model comparison including model probabilities of the second-level and free energy. (D) Parameter estimates for the flat and the hierarchical model. Error bars represent 95% credible intervals.

adaptation is considered as a means of the predictive brain to increase its precision, these findings can be integrated under the predictive coding framework (Garrido et al., 2009b). Thus, in accordance with the results of Parras and colleagues (2021, 2017), we found evidence for a hierarchical organization of prediction errors during deviance processing. Furthermore, our DCM analysis further underlined the hierarchical organization of the brain regions that indexed deviance processing. Thus, the increasing impact of predictive processing along the hierarchical organization does not only fit theoretical conceptualizations of the considered cortical nodes but is also empirically reflected in their causal relationships among each other.

While our study has many strengths, there are also limitations. In studies using electrophysiological measurements, SSA seems to be observed up to ISIs of 2 s (Ulanovsky et al., 2003). This is substantially shorter than the ISIs we chose here ($M = 3$ s). However, electrophysiological and hemodynamic evidence shows that in humans, deviance responses can still be observed using long ISIs of up to 10 s (Bottcher-Gandor and Ullsperger, 1992) or, on average, 4 s (Juckel et al., 2012). Furthermore, even in the animal model, adaptation phenomena can be observed for ISIs up to 60 s in some cortical regions (Netser et al., 2011). Besides, while it is undisputed that the ISI is an important factor influencing deviance responses (Netser et al., 2011; Ulanovsky et al., 2003;

Yarden and Nelken, 2017), the observation of differences between deviant and standard stimuli in a large number of different cortical regions in the current study speaks against the notion that deviance processing is completely abolished given ISIs > 2 s. Another difference between studies investigating SSA and our study is the irregularity of ISIs, which is often fixed in studies investigating SSA (Netser et al., 2011; Ulanovsky et al., 2003). This difference should be taken into account when comparing results. Taken together, the current study shows that an oddball sequence can be implemented in an event-related fMRI paradigm with comparably long ISIs. Future studies should try to find ways to implement shorter or fixed ISIs while still maintaining the advantages of event-related fMRI.

As a second limitation, the equiprobable control might not have been the ideal control condition to delineate adaptation effects as controls were not of equal probability compared to the standards in the oddball sequence. Another control condition like the cascade control might have solved this issue by providing a predictable control stimulus (Ruhnau et al., 2012). However, first, while different results might be expected from equiprobable and cascade control in theory, empirical evidence suggests that both are equally suited to investigate mechanisms of change detection (Parras et al., 2021, 2017; Wiens et al., 2019). Second, we were interested in how the contributions of several mechanisms might change across the auditory processing hierarchy. Importantly, even if the control condition was not optimal, changes between different cortices should still be meaningful. Third, while treated as equally predictable, the standard and the cascade control are not equal, as the next stimulus in the cascade control can be predicted with a probability of 100% while the standard in the oddball sequence cannot. Thus, while it is definitely desirable to use a cascade control in future studies, we think that in the current study, the use of the equiprobable control still allows us to draw conclusions on different mechanisms of deviance processing.

A third limitation is that we only investigated mechanisms of deviance processing during a condition where oddball stimuli had to be attended to but were not targets. This is different from many studies investigating deviance processing during preattentive conditions (Näätänen et al., 2011). While we can only speculate, we would expect that a task manipulation that diverts attention away from the oddball stimuli would possibly elicit deviance responses in fewer regions with less prominent prediction errors. Following this line of reasoning, mechanisms predominantly observed during deviance processing might be modulated by task settings. Experimentally combining task manipulations with fMRI might help better understand which brain regions vary in their predominant mechanisms and which do not. Besides, studies on repetition suppression have shown that the interplay between stimulus expectations and repetitions is complex, indicating that the influence of expectations on RS is not stable but changes with specific stimuli and conditions (Cacciaglia et al., 2019; Feuerriegel et al., 2018; Kovács et al., 2013; Tang et al., 2018; Todorovic et al., 2011; Utzerath et al., 2017). Thus, including different manipulations, e.g., of familiarity or higher-level expectations could further help in providing a comprehensive picture of hierarchical predictive processing in the brain. Furthermore, we only investigated the auditory modality. Future studies should also include other sensory stimulations. Finally, prospective research would profit from other analytical approaches for the investigation of different mechanisms of deviance processing. For example, computational modeling could offer a more targeted way to investigate different takes of the predictive processing account (e.g., Lieder et al., 2013).

5. Conclusion

We observed deviance-related effects in a widespread network of different brain regions. The processes predominantly responsible for these effects varied depending on the hierarchical level of the brain region. We detected an increase in prediction error-related activity and a concurrent decrease of adaptation-related activity from lower to higher hierarchi-

cal areas. These results highlight hierarchical predictive processing in the human brain.

Data and code availability statement

Data and code are available on the Open Science Framework accessible via <https://osf.io/eufd6/>.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Credit authorship contribution statement

Insa Schlossmacher: Conceptualization, Investigation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **Jacky Dilly:** Formal analysis, Investigation, Writing – review & editing. **Ina Protmann:** Formal analysis, Investigation, Writing – review & editing. **David Hofmann:** Investigation, Software, Writing – review & editing. **Torge Dellert:** Methodology, Software, Writing – review & editing. **Marie-Luise Roth-Paysen:** Investigation, Writing – review & editing. **Robert Moeck:** Investigation, Writing – review & editing. **Maximilian Bruchmann:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Thomas Straube:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Acknowledgments

The authors declare no competing financial interests. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We acknowledge support from the Open Access Publication Fund of the University of Münster.

References

- Alexander, W.H., Brown, J.W., 2019. The role of the anterior cingulate cortex in prediction error and signaling surprise. *Top. Cogn. Sci.* 11, 119–135. doi:10.1111/tops.12307.
- 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. doi:10.1016/j.neuroimage.2015.11.030.
- Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. *NeuroImage* 38, 95–113. doi:10.1016/j.neuroimage.2007.07.007.
- Asplund, C., Todd, J., Snyder, A., Marois, R., 2010. A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. *Nat. Neurosci.* 13, 507–512. doi:10.1038/nn.2509.
- Auksztulewicz, R., Friston, K., 2016. Repetition suppression and its contextual determinants in predictive coding. *Cortex* 80, 125–140. doi:10.1016/j.cortex.2015.11.024.
- Barron, H.C., Garvert, M.M., Behrens, T.E.J., 2016. Repetition suppression: a means to index neural representations using BOLD? *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150355. doi:10.1098/rstb.2015.0355.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528. doi:10.1093/cercor/10.5.512.
- Binkofski, F.C., Klann, J., Caspers, S., 2016. On the neuroanatomy and functional role of the inferior parietal lobule and intraparietal sulcus. In: *Neurobiology of Language*. Elsevier, pp. 35–47. doi:10.1016/B978-0-12-407794-2.00004-3.
- Bottcher-Gandor, C., Ullsperger, P., 1992. Mismatch negativity in event-related potentials to auditory stimuli as a function of varying interstimulus interval. *Psychophysiology* 29, 546–550. doi:10.1111/j.1469-8986.1992.tb02028.x.
- Cacciaglia, R., Costa-Faidella, J., Zarnowicz, K., Grimm, S., Escera, C., 2019. Auditory predictions shape the neural responses to stimulus repetition and sensory change. *NeuroImage* 186, 200–210. doi:10.1016/j.neuroimage.2018.11.007.
- Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi:10.1017/S0140525X12000477.
- Cole, M.W., Reynolds, J.R., Power, J.D., Repovs, G., Anticevic, A., Braver, T.S., 2013. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 16, 1348–1355. doi:10.1038/nn.3470.

- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi:10.1038/nrn755.
- Corbin, N., Todd, N., Friston, K.J., Callaghan, M.F., 2018. Accurate modeling of temporal correlations in rapidly sampled fMRI time series. *Hum. Brain Mapp.* 39, 3884–3897. doi:10.1002/hbm.24218.
- Costa, S.D., Zwaag, W.V.D., Marques, J.P., Frackowiak, R.S.J., Clarke, S., Saenz, M., 2011. Human primary auditory cortex follows the shape of Heschl's Gyrus. *J. Neurosci.* 31, 14067–14075. doi:10.1523/JNEUROSCI.2000-11.2011.
- Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31, 968–980. doi:10.1016/j.neuroimage.2006.01.021.
- Dürschmid, S., Edwards, E., Reichert, C., Dewar, C., Hinrichs, H., Heinze, H.-J., Kirsch, H.E., Dalal, S.S., Deouell, L.Y., Knight, R.T., 2016. Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *Proc. Natl. Acad. Sci.* 113, 6755–6760. doi:10.1073/pnas.1525030113.
- Eckert, M.A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., Dubno, J.R., 2009. At the heart of the ventral attention system: the right anterior insula. *Hum. Brain Mapp.* 30, 2530–2541. doi:10.1002/hbm.20688.
- Eklund, A., Nichols, T.E., Knutsson, H., 2016. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci.* 113, 7900–7905. doi:10.1073/pnas.1602413113.
- Feuerriegel, D., Keage, H.A.D., Rossion, B., Quek, G.L., 2018. Immediate stimulus repetition abolishes stimulus expectation and surprise effects in fast periodic visual oddball designs. *Biol. Psychol.* 138, 110–125. doi:10.1016/j.biopsycho.2018.09.002.
- Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. *NeuroImage* 19, 1273–1302. doi:10.1016/S1053-8119(03)00202-7.
- 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, 936–944. doi:10.1016/j.neuroimage.2008.05.018.
- Garrido, M.I., Kilner, J.M., Kiebel, S.J., Stephan, K.E., Baldeweg, T., Friston, K.J., 2009a. Repetition suppression and plasticity in the human brain. *NeuroImage* 48, 269–279. doi:10.1016/j.neuroimage.2009.06.034.
- Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009b. The mismatch negativity: A review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463. doi:10.1016/j.clinph.2008.11.029.
- Geuter, S., Boll, S., Eippert, F., Büchel, C., 2017. Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula. *eLife* 6, e24770. doi:10.7554/eLife.24770.
- Goulden, N., Khusnulnisa, A., Davis, N.J., Bracewell, R.M., Bokde, A.L., McNulty, J.P., Mullins, P.G., 2014. The salience network is responsible for switching between the default mode network and the central executive network: replication from DCM. *NeuroImage* 99, 180–190. doi:10.1016/j.neuroimage.2014.05.052.
- Greve, D.N., 2009. Optseq2.
- Heilbron, M., Chait, M., 2018. Great expectations: is there evidence for predictive coding in auditory cortex? *Neuroscience* 389, 54–73. doi:10.1016/j.neuroscience.2017.07.061.
- Ishishita, Y., Kunii, N., Shimada, S., Ibayashi, K., Tada, M., Kirihara, K., Kawai, K., Uka, T., Kasai, K., Saito, N., 2019. Deviance detection is the dominant component of auditory contextual processing in the lateral superior temporal gyrus: a human ECoG study. *Hum. Brain Mapp.* 40, 1184–1194. doi:10.1002/hbm.24438.
- Jääskeläinen, I.P., Ahveninen, J., Bonmassar, G., Dale, A.M., Ilmoniemi, R.J., Levänen, 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. *Proc. Natl. Acad. Sci.* 101, 6809–6814. doi:10.1073/pnas.0303760101.
- Jeffreys, H., 1961. *Theory of Probability*. UK Oxford University Press, Oxford.
- Juckel, G., Karch, S., Kawohl, W., Kirsch, V., Jäger, L., Leicht, G., Lutz, J., Stammel, A., Pogarell, O., Ertl, M., Reiser, M., Hegerl, U., Möller, H.J., Mulert, C., 2012. Age effects on the P300 potential and the corresponding fMRI BOLD-signal. *NeuroImage* 60, 2027–2034. doi:10.1016/j.neuroimage.2012.02.019.
- Kastner, S., Ungerleider, L.G., 2000. Mechanisms of visual attention in the human cortex. *Kiehl, K.A., Stevens, M.C., Laurens, K.R., Pearlson, G., Calhoun, V.D., Liddle, P.F., 2005. An adaptive reflexive processing model of neurocognitive function: supporting evidence from a large scale (n = 100) fMRI study of an auditory oddball task. NeuroImage 25, 899–915. doi:10.1016/j.neuroimage.2004.12.035.*
- Kim, H., 2014. Involvement of the dorsal and ventral attention networks in oddball stimulus processing: a meta-analysis. *Hum. Brain Mapp.* 35, 2265–2284. doi:10.1002/hbm.22326.
- Kovács, G., Kaiser, D., Kaliukhovich, D.A., Vidnyánszky, Z., Vogels, R., 2013. Repetition probability does not affect fMRI repetition suppression for objects. *J. Neurosci.* 33, 9805–9812. doi:10.1523/JNEUROSCI.3423-12.2013.
- Larsson, J., Solomon, S.G., Kohn, A., 2016. fMRI adaptation revisited. *Cortex* 80, 154–160. doi:10.1016/j.cortex.2015.10.026.
- Laufer, I., Negishi, M., Rajeevan, N., Lacadie, C.M., Constable, R.T., 2008. Sensory and cognitive mechanisms of change detection in the context of speech. *Brain Struct. Funct.* 212, 427–442. doi:10.1007/s00429-007-0167-8.
- Li, Q., Liu, G., Yuan, G., Wu, Z., Zhao, X., Zhao, Xingcong, 2019. Single-trial EEG-fMRI reveals the generation process of the mismatch negativity. *Front. Hum. Neurosci.* 13. doi:10.3389/fnhum.2019.00168.
- Lieder, F., Daunizeau, J., Garrido, M.I., Friston, K.J., Stephan, K.E., 2013. Modelling trial-by-trial changes in the mismatch negativity. *PLoS Comput. Biol.* 9, e1002911. doi:10.1371/journal.pcbi.1002911.
- Maess, B., Jacobsen, T., Schröger, E., Friederici, A.D., 2007. Localizing pre-attentive auditory memory-based comparison: magnetic mismatch negativity to pitch change. *NeuroImage* 37, 561–571. doi:10.1016/j.neuroimage.2007.05.040.
- May, P.J.C., Tiitinen, H., 2010. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47, 66–122. doi:10.1111/j.1469-8986.2009.00856.x.
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667. doi:10.1007/s00429-010-0262-0.
- Näätänen, R., Kujala, T., Winkler, I., 2011. Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology* 48, 4–22. doi:10.1111/j.1469-8986.2010.01114.x.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi:10.1016/j.clinph.2007.04.026.
- Nelken, I., 2014. Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biol. Cybern.* 108, 655–663. doi:10.1007/s00422-014-0585-7.
- Netser, S., Zahar, Y., Gutfreund, Y., 2011. Stimulus-specific adaptation: can it be a neural correlate of behavioral habituation? *J. Neurosci.* 31, 17811–17820. doi:10.1523/JNEUROSCI.4790-11.2011.
- O'Connor, A.R., Han, S., Dobbins, I.G., 2010. The inferior parietal lobule and recognition memory: expectancy violation or successful retrieval? *J. Neurosci.* 30, 2924–2934. doi:10.1523/JNEUROSCI.4225-09.2010.
- Olshowy, W., Aston, J., Rua, C., Williams, G.B., 2019. Accurate autocorrelation modeling substantially improves fMRI reliability. *Nat. Commun.* 10, 1220. doi:10.1038/s41467-019-09230-w.
- Opitz, B., Schröger, E., Cramon, D.Y.V., 2005. Sensory and cognitive mechanisms for preattentive change detection in auditory cortex. *Eur. J. Neurosci.* 21, 531–535. doi:10.1111/j.1460-9568.2005.03839.x.
- O'Shea, R.P., 2015. Refractoriness about adaptation. *Front. Hum. Neurosci.* 9. doi:10.3389/fnhum.2015.00038.
- Parras, G.G., Casado-Román, L., Schröger, E., Malmierca, M.S., 2021. The posterior auditory field is the chief generator of prediction error signals in the auditory cortex. *NeuroImage* 242, 118446. doi:10.1016/j.neuroimage.2021.118446.
- Parras, G.G., Nieto-Diego, J., Carbajal, G.V., Valdés-Baizabal, C., Escera, C., Malmierca, M.S., 2017. Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nat. Commun.* 8, 2148. doi:10.1038/s41467-017-02038-6.
- Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi:10.1016/j.clinph.2007.04.019.
- Ruhnau, P., Herrmann, B., Schröger, E., 2012. Finding the right control: the mismatch negativity under investigation. *Clin. Neurophysiol.* 123, 507–512. doi:10.1016/j.clinph.2011.07.035.
- Schröger, E., Wolff, C., 1996. Mismatch response of the human brain to changes in sound location. *NeuroReport* 7, 3005–3008. doi:10.1097/00001756-199611250-00041.
- Sestieri, C., Shulman, G.L., Corbetta, M., 2012. Orienting to the environment: separate contributions of dorsal and ventral frontoparietal attention networks. In: Mangun, G.R. (Ed.), *The Neuroscience of Attention: Attentional Control and Selection*. Oxford University Press, New York, pp. 100–130. doi:10.1093/acprof/oso/9780195334364.003.0005.
- Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., Davidson, R.J., 2011. The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* 12, 154–167. doi:10.1038/nrn2994.
- Siman-Tov, T., Granot, R.Y., Shany, O., Singer, N., Hendler, T., Gordon, C.R., 2019. Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. *Neurosci. Biobehav. Rev.* 105, 262–275. doi:10.1016/j.neubiorev.2019.08.012.
- Stefanics, G., Kremláček, J., Czigler, I., 2014. Visual mismatch negativity: a predictive coding view. *Front. Hum. Neurosci.* 8, 666. doi:10.3389/fnhum.2014.00666.
- Tang, M.F., Smout, C.A., Arabzadeh, E., Mattingley, J.B., 2018. Prediction error and repetition suppression have distinct effects on neural representations of visual information. *eLife* 7, e33123. doi:10.7554/eLife.33123.
- Todorovic, A., van Ede, F., Maris, E., de Lange, F.P., 2011. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. *J. Neurosci.* 31, 9118–9123. doi:10.1523/JNEUROSCI.1425-11.2011.
- Ulanovsky, N., Las, L., Nelken, I., 2003. Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* 6, 391–398. doi:10.1038/nrn1032.
- Utzerath, C., St. John-Saaltink, E., Buitelaar, J., de Lange, F.P., 2017. Repetition suppression to objects is modulated by stimulus-specific expectations. *Sci. Rep.* 7, 8781. doi:10.1038/s41598-017-09374-z.
- Wiens, S., Szychowska, M., Eklund, R., van Berlekom, E., 2019. Cascade and no-repetition rules are comparable controls for the auditory frequency mismatch negativity in oddball tasks. *Psychophysiology* 56, e13280. doi:10.1111/psyp.13280.
- Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., Nichols, T.E., 2014. Permutation inference for the general linear model. *NeuroImage* 92, 381–397. doi:10.1016/j.neuroimage.2014.01.060.
- 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. *Int. J. Psychophysiol.* 83, 132–143. doi:10.1016/j.ijpsycho.2011.10.001.
- Yan, C.-G., Wang, X.-D., Zuo, X.-N., Zang, Y.-F., 2016. DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* 14, 339–351. doi:10.1007/s12021-016-9299-4.
- Yarden, T.S., Nelken, I., 2017. Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLOS Comput. Biol.* 13, e1005437. doi:10.1371/journal.pcbi.1005437.

Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi:[10.1152/jn.00338.2011](https://doi.org/10.1152/jn.00338.2011).

Zeidman, P., Jafarian, A., Corbin, N., Seghier, M.L., Razi, A., Price, C.J., Friston, K.J., 2019a. A guide to group effective connectivity analysis, part 1: first level analysis

with DCM for fMRI. *NeuroImage* 200, 174–190. doi:[10.1016/j.neuroimage.2019.06.031](https://doi.org/10.1016/j.neuroimage.2019.06.031).

Zeidman, P., Jafarian, A., Seghier, M.L., Litvak, V., Cagnan, H., Price, C.J., Friston, K.J., 2019b. A guide to group effective connectivity analysis, part 2: second level analysis with PEB. *NeuroImage* 200, 12–25. doi:[10.1016/j.neuroimage.2019.06.032](https://doi.org/10.1016/j.neuroimage.2019.06.032).