Hierarchical modulation of auditory prediction error signaling is independent of attention

Kristiina Kompus, Vegard Volehaugen, Juanita Todd & René Westerhausen

View supplementary material 🗗

Published online: 01 Aug 2019.

Submit your article to this journal 🗗

View Crossmark data 🗗

Hierarchical modulation of auditory prediction error signaling is independent of attention

Kristiina Kompus[a,b], Vegard Volehaugen[a], Juanita Todd[c] and René Westerhausen[a,d]

[a]Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway; [b]Institute of Psychology, University of Tartu, Tartu, Estonia; [c]School of Psychology, University of Newcastle, Newcastle, Australia; [d]Department of Psychology, University of Oslo, Oslo, Norway

ABSTRACT

The auditory system is tuned to detect rhythmic regularities in the environment which can occur on different timescales. Event-related potentials such as mismatch negativity (MMN) and P3b are thought to index local and global deviance, respectively. However, it is not clear how these hierarchical levels interact and to what extent attention modulates this interaction. In this EEG study with 17 healthy young adults, we used a hierarchical oddball paradigm with local (sequence-level) and global (block-level) violations in attended and unattended conditions. Amplitude of N2 and P3b were analyzed in a 2*2*2 factorial model (local status, global status, attention condition). We found a significant interaction between the local and global status on the N2 amplitude, while there was no significant three-way interaction with attention, together demonstrating that lower-level prediction error is modulated by detection of higher-order regularity but expressed independently of attention. By contrast, higher-level prediction error, indexed by P3b, was sensitive to global regularity violations if the auditory stream was attended. The results demonstrate the capacity of our auditory perception to preattentively resolve conflicts between different levels of predictive hierarchy even across longer time intervals as indexed by MMN modulation, while P3b represents a different, attention-dependent system.

1. Introduction

Our perception relies on prediction to facilitate the decoding of the sensory information. The predictive coding theories of perception suggest that the brain tries to minimize the surprise or prediction error, and continuously uses the unpredicted portion of the sensory input to adjust the predictive models (Friston, 2005). A crucial component of the predictive coding is the hierarchical organization of perceptual systems, with higher levels which represent slower-changing regularities modulating the processing of lower-level predictive units which integrate over a shorter time (Kiebel, Daunizeau, & Friston, 2008). Such nested hierarchical system is crucial in human speech processing, where the probability of a sound depends on its immediate local environment such as the syllable structure, whereas word and sentence rules in the given language give wider, global context which would need to be taken into account when predicting the subsequent sound (Hickok, Houde, & Rong, 2011). How such hierarchically nested local and global rules are extracted from the auditory stream and how they interact with each other as well as other systems such as attention and long-term memory are central questions to our understanding of auditory perception.

It has been suggested that simple tone *vs.* complex pattern deviations, corresponding to violating local and global rules, respectively, are processed by different neural systems, with simple-feature deviance-detection occurring at earlier levels of auditory processing and increasingly complex rule deviations detected on higher levels (Cornella, Leung, Grimm, & Escera, 2012). Crucially, global-rule formation has been suggested to be attention-dependent. An influential series of studies (Chennu et al., 2013; Bekinschtein et al., 2009; Marti, Thibault, & Dehaene, 2014; Strauss et al., 2015) has used a local/global paradigm where five-tone sequences (AAAAA or AAAAB) are presented within a longer block of sequences, where each sequence can be either typical to the block (i.e., majority of the sequences are the same type) or differ from the rule imposed by the majority within the block. In this situation two hierarchically distinct predictions can be made: local predictions about the fifth tone in the sequence, and global predictions about the probability

CONTACT Kristiina Kompus ✉ kko025@uib.no 🖃 Department of Biological and Medical Psychology, University of Bergen, Jonas Lies vei 91, Bergen 5009, Norway

of the fifth tone while taking into account the probability of the five-tone sequence in the context of the current block. These studies have suggested that there is an interaction between the global deviance processing and attention: while the local rule representation is independent of attention, detecting the deviance from the global rules is attention-dependent. With this paradigm, the violation of the global rules (where the five-tone sequence differs from a rule of the group) has been found to elicit P3b component, a parietal positive wave at 300–600 ms which is not elicited when the subjects are not actively attending the stimuli. This has led to the suggestion that the global-rule representation is selectively performed in neural structures which are indexed by the P3b generation, and the operation of these structures is reliant on the availability of the global workspace (Bekinschtein et al., 2009). After this attention-dependent higher-order rule has been formed, it can regulate the 'local' effect which is elicited in the mismatch negativity (MMN) timeframe, 150–250 ms after the onset of the deviant (Wacongne et al., 2011). Such dual-process view with attention-dependent context processing, however, is not in agreement with studies that have shown pattern-violation MMN for longer patterns even in the absence of conscious awareness (see Paavilainen, 2013 for overview). For example, Herholz and colleagues (Herholz, Lappe, & Pantev, 2009) found that a pattern-violation MMN was generated by a deviation from a four-tone pattern even when the subjects did not consciously notice the pattern (see also Horvath, Czigler, Sussman, & Winkler, 2001). Also, complex contingency rules on the predictability of the sound features have been shown to modulate MMN without attention to the stimuli (Todd & Robinson, 2010).

In this study, we used the local/global paradigm under attended and unattended condition to examine whether the interaction between the local and global predictions is modulated by whether or not the stimuli are attended. In case the interaction between the local and global rules in the MMN timeframe would differ between an attended and unattended task situation, then the extraction of a global regularity from an auditory stream was indeed attention-dependent. By contrast, in case attention only affects the P3b elicitation but not the local–global interaction for the MMN peak expressed in the N2 timeframe, it would suggest that the global-rule extraction is performed also without attention, and the P3b represents a different process which is attention-dependent, but independent of the global-rule representation in the auditory stimulus processing.

In addition to time windows corresponding to the N2 and P3b, we also examined the effect of the local and global-rule violation under attended and unattended conditions on the P3a, a component with a frontocentral maximum following the negativity in the MMN timeframe, associated with the attention capture by the deviant stimulus. The P3a has been suggested to index the attentional capture by the deviant stimulus (Escera & Corral, 2007), and appears to be generated when the MMN reaches a certain threshold value. If the P3a demonstrates the interaction of the local and global predictions, this would show that the attentional capture-related response is either directly coupled to the magnitude of the MMN, or sensitive to the same contextual modulation as the MMN. By contrast, a lack of interaction would support the view that P3a generation is not directly coupled to the process that creates MMN.

We expect that the MMN amplitude will depend on both the local and global status of the tone, that the P3b is generated only for global deviants in the attended condition, and that the attention condition does not affect the MMN amplitude modulation by local/global status.

2. Methods

2.1. Participants

The participants were 20 healthy young adults (10 females) with no history of psychiatric or neurological disease by self-report and auditory thresholds <25 dB in both ears for frequencies 250, 500, 1000, 2000 and 3000 Hz as assessed using the Hughson-Westlake audiometric test (Oscilla USB-300, Inmedico, Lystrup, Denmark). Three subjects' data were discarded due to excess movement artifacts during EEG session; the analyses are based on 17 subjects (7 females, mean age 25.4, SD 3.6). The subjects were asked to refrain from nicotine and caffeine for at least 5 h before the recording. The participants were informed about the experimental procedures and signed an informed consent form. The study was approved by the regional ethics committee (REK-VEST).

2.2. Paradigm

The 'Local–Global' MMN paradigm consisted of sequences of five harmonic tones composed of three sinusoidal partials (tone A: 500, 1000, 1500 Hz; tone B: 550, 1100, 1600 Hz) with 50-ms duration (including 7 ms rise and fall time). Interval between tones in a sequence was 100 ms, with total sequence length 650 ms. The sequences were either XXXXX (five identical tones, AAAAA and BBBBB) or XXXXY (fifth tone different, AAAAB and BBBBA). The final

tone in the sequence could thus be either local standard or local deviant.

The sequences were presented in blocks, with inter-sequence interval 700–900 ms (random variation with step 50 ms, mean 800 ms). Half of the blocks had 80% XXXXX sequences and 20% XXXXY sequences, in the other half the frequency was reversed (80% XXXXY sequences, 20% XXXXX sequences). The sequences could thus be global standards (the dominant sequences) or global deviants (the rare sequences). This leads to a factorial design to probe the final tone of the sequence (global status: deviant/standard, local status: deviant/standard).

Each block began with 25 repetitions of the sequence which was the global standard for that block, serving as a model-building phase, following which 150 sequences were presented (in 80/20 ratio). The block length was 4 min 22 s. Inter-block interval was 3 s. In total, eight blocks were presented: four with XXXXX as global standard and four with XXXXY as global standard. In total, there were 480 trials in the global standard-local deviant as well as global standard-local standard conditions, and 120 trials in the global deviant-local deviant as well as global deviant-local standard condition.

Two attentional conditions were used, with order counterbalanced across subjects: attended and unattended stimulation. In the attended condition the subjects were asked to monitor the tones, keeping eyes on a fixation cross in the center of a monitor. Compliance in the attended condition was checked by asking the subjects to report on the sound characteristics after the recording using a 5-item questionnaire similar to Bekinschtein et al. (2009) regarding awareness of regularities in the sounds and presence of irregular sequences of each type. All subjects reported detecting some of the regularities present in the sound streams, with the average score of 4.1 out of 5 possible (only three subjects scoring 2 or 3, the others 4 or 5). In the unattended condition, the subjects were asked to ignore the sounds and concentrate on a visual working memory task. In the visual n-back task abstract visual objects (Fribbles, TarrLab, http://www.tarrlab.org) were presented, asynchronously with the auditory stimuli, and the subjects asked to press a button in case the object was the same as two objects previously. Compliance was checked by examining the response profile of the visual n-back task.

2.3. EEG recording and analysis

The subjects were presented with the tones through headphones while data were collected in an electromagnetically shielded EEG recording chamber using 12 Ag/AgCl electrodes (F3, Fz, F4, FCz, C3, Cz, C4, TP9, TP10, P3, Pz, P4) placed according to the International 10–20 system using the EasyCap electrode cap (Falk Minow Services, Breitenbrunn, Germany) and Abralyt 2000 electrode gel. Interelectrode impedance was kept under 10 kΩ. The reference electrode was placed at nosetip, the ground at FT10. Four electrodes were used for monitoring eye movements, two placed above and below the right eye and at the outer canthi of the eyes. The EEG data were recorded using BrainVision Recorder 1.0 (Brain Products, Munich, Germany). The sampling rate was 500 Hz, filter band 0–100 Hz. The subjects were instructed prior to the start of recording to avoid movement and excess eye movements and blinks.

After recording, the data were offline filtered using a zero-phase Butterworth IIR filter with high-pass threshold 0.01 Hz (slope 12 dB/oct) and low-pass threshold 30 Hz (slope 12 dB/oct). The data were down-sampled to 250 Hz. Eye movements were removed using Gratton–Coles algorithm implemented in the BrainVision Analyzer. The data were epoched into segments relative to the onset of the first tone in the 5-tone sequence. Epochs spanned from −100 to 1348 ms after the onset of the first tone, covering the entire 5-tone sequence. Automatic artifact detection was used, with epochs where the amplitude exceeded +-100 μV were discarded. On average 17% of the trials were discarded, with no significant difference in the proportion of discarded trials between the conditions ($F(7,128) = 0.45$, $p = .87$). The final tone (which could deviate from the local or global rule) onset was at 650 ms.

To reduce the effect of contingent negative variation between attended and unattended conditions, which could have affected the waveform from the beginning of the first tone, the epochs were corrected to the mean of the 50-ms baseline before the onset of the final tone of the sequence (which could be physical deviant or standard). To ensure that our results do not depend on the choice of the baseline and are directly comparable with previous studies using the local-global paradigm where various choices for the baseline correction were used (cf. Chennu et al., 2013; Bekinschtein et al., 2009; Wacongne et al., 2011), we also performed analyses using correction to baseline over the time window −100 to 0 ms before the onset of the first tone in the five-tone sequence as well as peak-to-peak analyses; these results are reported in Supplementary Results. Relative to the onset of the final tone, the N2 was quantified as the most negative value (averaging over ± 20 ms around the peak) in the time window 50–250 ms in the electrode FCz; the mastoid positivity was examined as the most positive value in the same time window at the electrodes TP9 and TP10 (Duncan et al., 2009) to identify whether the frontocentral negativity

corresponds to MMN by showing simultaneous positivity in the mastoid locations (Duncan et al., 2009). We tested the N2 peak in the raw average waveform, for comparability with the previous studies using the local-global paradigm as used here. However, to reduce the contribution of other, attention-dependent potentials in the same timeframe, we also calculated the difference waves between the standard and global sequences and examined the MMN as a negativity evident in the difference wave between the three types of deviant stimuli (local-only, global-only, local and global) and the standard stimulus (local and global standard). The three difference waves in both attended and unattended conditions were calculated, and for each subject, the most negative value in the time window 50–250 ms was extracted (mean around ± 20 ms surrounding the peak).

The P3a was quantified as the most positive value following N2 up to 400 ms post-deviant onset at the electrode FCz (Polich, 2007). Finally, the P3b was quantified as the area under the curve in the interval 300–500 ms at the electrode Pz (Polich, 2007).

2.4. Statistical analysis

Extracted peak value analysis was performed on IBM SPSS Statistics version 25. The data were analyzed using a 2*2*2 repeated-measures general linear model (GLM), with factors Local status (standard/deviant), Global status (standard/deviant) and Attention (attended/unattended). Greenhouse-Geisser correction was applied where appropriate. The effect sizes were calculated as eta squares (η^2), for post-hoc paired comparison the Cohen's d was calculated. The effects of interest are the two-way interaction between Local and Global factors, and the tree-way interaction between Local, Global and Attention for the peak values in the time window 50–250 ms (N2) as well as P3a time window, testing the contextual modulation of local deviance detection by N2 and P3a, respectively, and testing whether the contextual modulation of these peaks is affected by attention. In addition, in the P3b time window, the effect of interest is the main effect of Global and interaction between Global and Attention, testing the sensitivity to the global rule as well as the attention-dependence of global-rule representation by the P3b. The significance level α was set at 0.05 for all of the four analyses. Correction for multiple analyses was not implemented to secure statistical power also for smaller population effects and reduce the likelihood of false-negative results. To support interpretation, we also conducted a sensitivity power analysis in G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009; www.gpower.hhu.de/), that is, the population effects size (expressed as proportion explained variance) that can be reliably excluded with a test power of .80 given the present sample size and design.

In addition to the windowed analysis in preselected time intervals, we also estimated the GLM as specified above (including a subject factor to account for the within-subjects design) at each timepoint in each of the electrodes after the onset of the final tone to examine the time course of the attentional effects on the local and global status and their interaction on the ERPs. For significance testing we used the PALM package (Winkler, Ridgway, Webster, Smith, & Nichols, 2014) with the threshold-free cluster enhancement procedure, performing 8000 permutations; the p-values were corrected using familywise error correction across all timepoints.

3. Results

During the attention manipulation, the performance indicators showed that all subjects gave responses to the visual task, indicating compliance with task instruction. The mean accuracy (proportion of hits and correct rejections) was 0.87 (range 0.77–0.95, sd 0.06); the mean sensitivity index d' was 1.65 (range 0.82–2.74, sd 0.62).

For the EEG data, we tested the amplitude for N2 and MMN as well as P3a and P3b, examining for the effect of Global, Local and Attention factors and all interactions.

3.1. N2 and MMN

The waveforms for N2 as well as difference waveforms between the standard and deviant ERPs are depicted in Figure 1, row A and B. As can be seen in the difference waveforms (Figure1(b)), a clear MMN was present in all conditions; testing the peak amplitude at FCz electrode against zero using a one-sample t-test indicated that the peak value was significantly below zero in all six conditions (all p < .005). We then examined the effect of the different experimental conditions on the N2 peak (Figure1(a)) using the factorial GLM as described above. In FCz, there was a main effect of Local status (F(1,16) = 22.4; p < .001, η^2 = .21), Global status (F(1,16) = 87.1; p < .001, η^2 = .35) and Local*Global interaction (F(1,16) = 19.6; p < .001, η^2 = .09). There were no significant main effects or interactions involving the Attention factor. The three-way interaction was non-significant (F(1,16) = 0.12; p = .74) showing an empirical effect size of η^2 < .001 which can be considered negligible. Sensitivity power analysis further suggests that population effects larger than 6% explained variance can be excluded. Post-hoc pairwise comparisons, collapsing over the
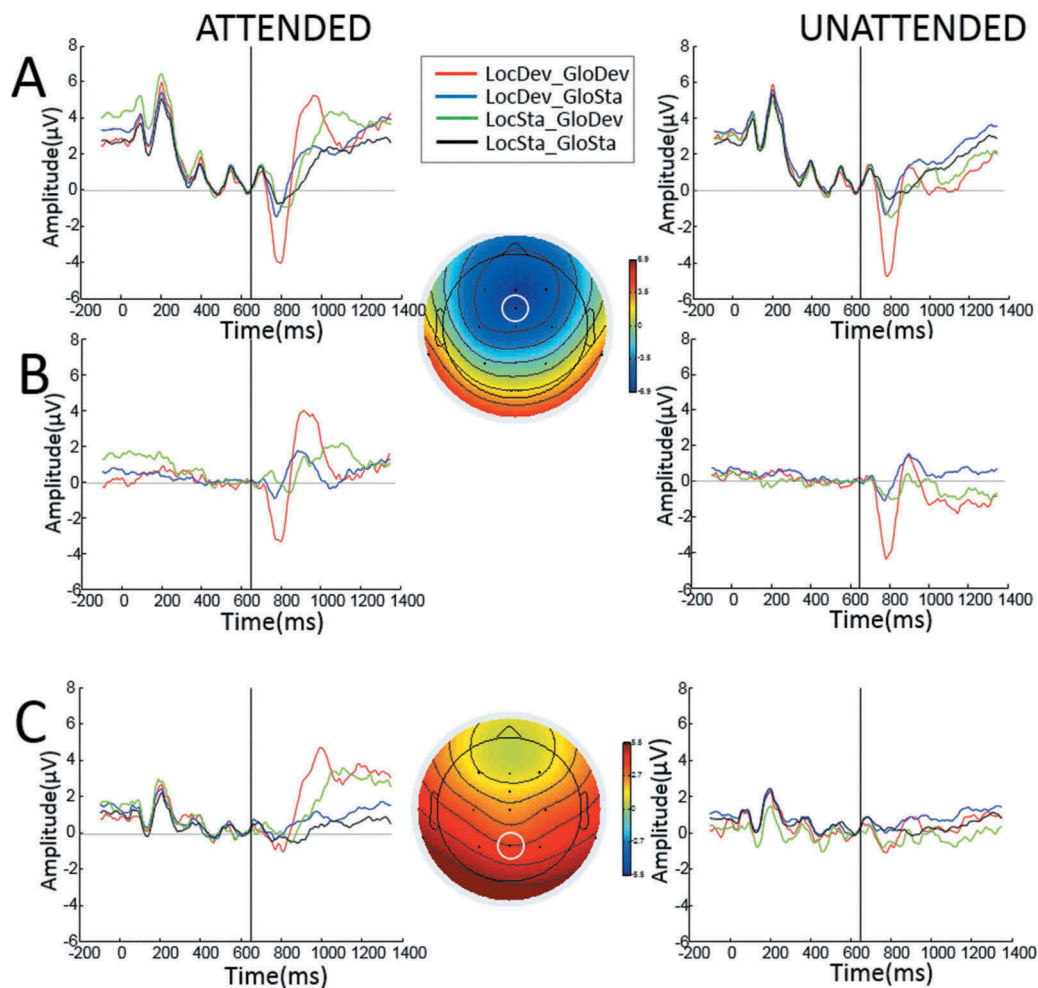
Figure 1. (a): The waveforms from FCz (marked on the upper topographic plot showing the distribution of voltage) depicting the N2 and the P3a for attended (left) and unattended (right) condition. (b): The difference waveforms from FCz for the three deviant condition relative to the local and global standard condition, isolating the MMN component. All MMN peaks are significantly below zero (see text). The GLM with factors Attention and Type in FCz showed a significant main effect of Type (F(2,32) = 33.2, p < .001), with no significant main or interaction effects of Attention. Pairwise comparison between the three types showed a significant difference between local-global and the other two types. (c): The waveforms from Pz (marked on the lower topographic plot showing the distribution of voltage) depicting the P3b for attended (left) and unattended (right) condition.

Attention factor, showed that the standard *vs* deviant effect for the Local factor depended on the level of the Global factor: for items that were also global standards, the local standard-deviant difference was smaller (sta: −.93, dev: −1.45; t (16) = −1.6, p = .12, d = 0.36) than for items that were also global deviants (sta: −1.9, dev:-4.5; t(16) = −5.5, p < .001, d = 1.5). We also examined the effect of the global status on the local standards, comparing the items that were local standards-global deviants and local standards-global standards directly. There was a significant difference between the local standard items that were global standards vs global deviants (global sta: −.93, global dev: −1.9; t (16) = −4.4, p < .001, d = 1.04). Thus, the final X-tone in the sequence, a stimulus which was physically identical to the

preceding tones and had a very high probability overall, elicited a deviance response due to the higher-order rule governing the occurrence of the physically deviating Y-tones. This highlights that the global rule was represented in the N2 time window as it led to increased negativity in the absence of a local deviant.

For the mastoid electrodes (Figure 2), additionally, the Laterality factor was included in the GLM to assess whether the effects differ between left and right side. As there was a three-way interaction between Laterality, Local and Global (F(1,16) = 14.0, p = .002, η² = .01), we performed separate GLMs in each of the electrode sites. In the left mastoid, there was a significant main effect of Local status (F(1,16) = 18.9, p = .001, η² = .19), Global status (F(1,16) =
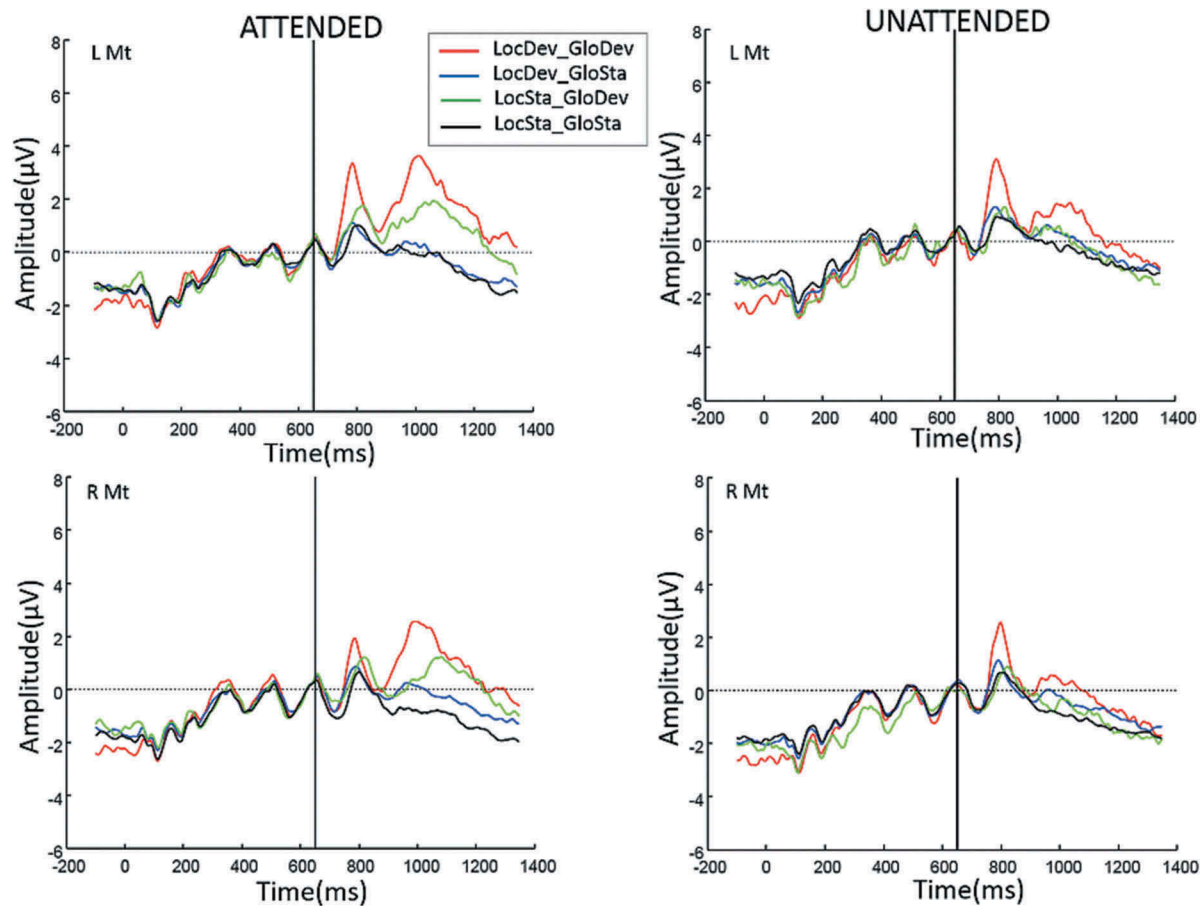
Figure 2. The waveforms from left and right mastoid (top: L Mt; bottom: R Mt) depicting the mastoid positivity for attended (left) and unattended (right) condition.

58.5, $p < .001$, $\eta^2 = .46$) and Local*Global interaction (F(1,16) = 29.3, $p < .001$, $\eta^2 = .11$). Same pattern was found in the right mastoid: there was a significant main effect of Local status (F(1,16) = 10.5, $p = .005$, $\eta^2 = .14$), Global status (F(1,16) = 13.8, $p = .002$, $\eta^2 = .24$) and Local*Global interaction (F(1,16) = 7.7, $p = .014$, $\eta^2 = .03$). There were no significant main effects or interactions involving the Attention factor in either of the mastoid electrodes. The three-way interaction between the three factors (Local*Global*Attention) was not significant and had a very small effect size both in the left (F(1,16) = 0.01; $p = .92$; $\eta^2 < .001$) and right mastoid (F(1,16) = 0.89; $p = .36$; $\eta^2 = .01$). As for the N2 analysis, population effects larger than 6% explained variance can be excluded. The effect sizes suggest that the three-way interaction involving the electrode factor was due to the difference in effect sizes (left > right), but not due to differences in the overall pattern of associations. Post-hoc pairwise comparisons, collapsing over Attention factor, showed the same pattern as in FCz: the standard vs deviant effect for the Local factor depended on the level of the Global factor. In the left mastoid, for items that were also global standards, the local standard-deviant difference was smaller (sta: 1.1, dev 1.3, $t = 1.11$, $p = .28$, $d = 0.28$) than for the items that were also global deviants (sta: 1.8, dev: 3.34, $t = 5.5$, $p < .001$, $d = 1.3$). In the right mastoid, a similar pattern was seen: for items that were also global standards, the local standard-deviant difference was smaller (sta: .66, dev: 1.03, $t = 2.09$, $p = .05$, $d = .53$) than for the items that were also global deviants (sta: 1.3, dev 2.28, $t = 3.4$, $p = .004$, $d = .91$). Comparing the local standards depending on the global status showed that there was a significant difference between the local standard items that were global standards vs global deviants in both the left mastoid (global sta: 1.11, global dev: 1.8; $t(16) = 3.32$, $p = .004$, $d = .81$) as well as the right mastoid (global sta: .66, global dev: 1.25; $t(16) = -4.4$, $p = .04$, $d = .54$). This highlights that the global rule was represented in the mastoids similarly as in the fronto-central electrode.

3.2. P3a

The waveform demonstrating P3a is depicted in Figure 1(a). There was a main effect of Local status (F(1,16) = 6.6, p = .02, η² = .02) and Global status (F(1,16) = 14.5, p = .002 η² = .11), with deviant stimuli eliciting larger amplitude than standard stimuli, and Attention (F(1,16) = 10.9, p = .005 η$_p$² = .23), with attended stimuli eliciting larger amplitude than unattended stimuli. The Local*Global interaction as well as Local*Global*Attention interaction were not significant. However, there was a significant interaction between Attention*Global (F(1,16) = 16.7, p = .001 η² = .09). Post-hoc pairwise comparisons, collapsing over the Local factor showed that the global-deviants and global-standards differed only in the attended condition (5.58 vs 2.7, t = 4.6, p < .001, d = 1.6); while there was no significant difference in the unattended condition (1.99 vs 1.83, t = .42, p = .68, d = 0.11).

3.3. P3b

The waveform showing P3b in the electrode Pz is depicted on Figure1(c). There was a significant main effect of Global status (F(1,16) = 11.6, p = .004; η² = .08), Attention (F(1,16) = 13.8, p = .002; η² = .27), and a significant interaction between Global status and Attention (F(1,16) = 38.0, p < .001; η² = .22). Post-hoc pairwise comparisons, collapsing over the Local factor showed that the global-deviants compared to global-standards were more positive only during attended condition (3.2 vs .68, t = 5.4, p < .001, d = 2.2), whereas during unattended condition the deviants were slightly more negative than standards (−.07 vs .52, t = −2.2, p = .04, d = −0.58). Other effects were not significant.

3.4. Timepoint-by-timepoint analysis

The results of the permutation tests at each timepoint after the onset of the final tone to examine the time course of the attentional effects on the local and global status and their interaction are shown in Figure 3. They indicate that the Local*Global interaction effect demonstrated in the peak analysis above extends across the frontal and central electrodes, stronger on the midline and right, whereas the Attention*Global interaction shown for P3b encompasses not only the parietal but also central electrodes and lasts from approximately 300-ms post-deviant onset to the end of the epoch. The timepoint-by-timepoint analysis also confirms that the three-way interaction between local and global status and attention is not significant anywhere, at most a weak trend can be seen in a small number of

timepoints in electrodes F3 and C3. Also, attention does not significantly interact with the Local factor anywhere, supporting the notion that MMN generation is not modulated by attention, at least in situations where attention does not lead to substantial reorganization of the auditory stream. Of note is a discrepancy between Local and Global factors in the P3a: while the Local deviants lead to a significant effect on the P3a range starting quite early (~200 ms post-deviant and centrally located, suggesting that it is the early, attention-independent phase of the P3a), by contrast, the effect did not survive the familywise error correction for the Global factor. The later, 300-ms onset effect in frontocentral and central electrodes visible in Attention*Global plot could be related to the later phase of the P3a. Finally, it is noticeable that both Local and Global factors have a significant effect on the bilateral mastoid electrodes after ~250 ms post-deviant onset, and the effect extends longer in the Global than Local.

4. Discussion

In this experiment, we exposed subjects to a hierarchical auditory structure where the frequency of a tone was predicted by two independent rules in a factorial design: a local rule in relation to the immediate environment and a global rule applying over a longer timeframe. The local rule violation as well as the global rule violation elicited negativity in the N2 timeframe, having the characteristics of a classical MMN response (frontal negativity and simultaneous mastoid positivity). This demonstrates that both local and global rules were represented in the N2 time window, contrary to studies which have suggested that global-rule violation is indexed by P3b peak (Bekinschtein et al., 2009; Chennu et al., 2013; Cornella et al., 2012; Wacongne et al., 2011). Further, the effect of the independently varying global and local rule on predicting the same feature (frequency) is interactive: the amplitude of the N2 is the product of the tone's status on both levels. This expands the current view on the representation of multiple rules simultaneously: while it has been previously shown that the different features of a tone are represented independently and that the resulting MMN is additive (Takegata, Paavilainen, Näätänen, & Winkler, 1999), the current findings emphasize that in the case of hierarchically dependent rules about the same feature, the auditory system calculates the 'joint surprise' based on an interaction of these rules. Finally, we show that the global-rule representation is not dependent on attention, as the N2 amplitude increase for deviant stimuli (local, global and local*global) could be seen in both attended and unattended conditions; and attention did not significantly affect the magnitude of the amplitude change. While our results disagree with the suggestions from the
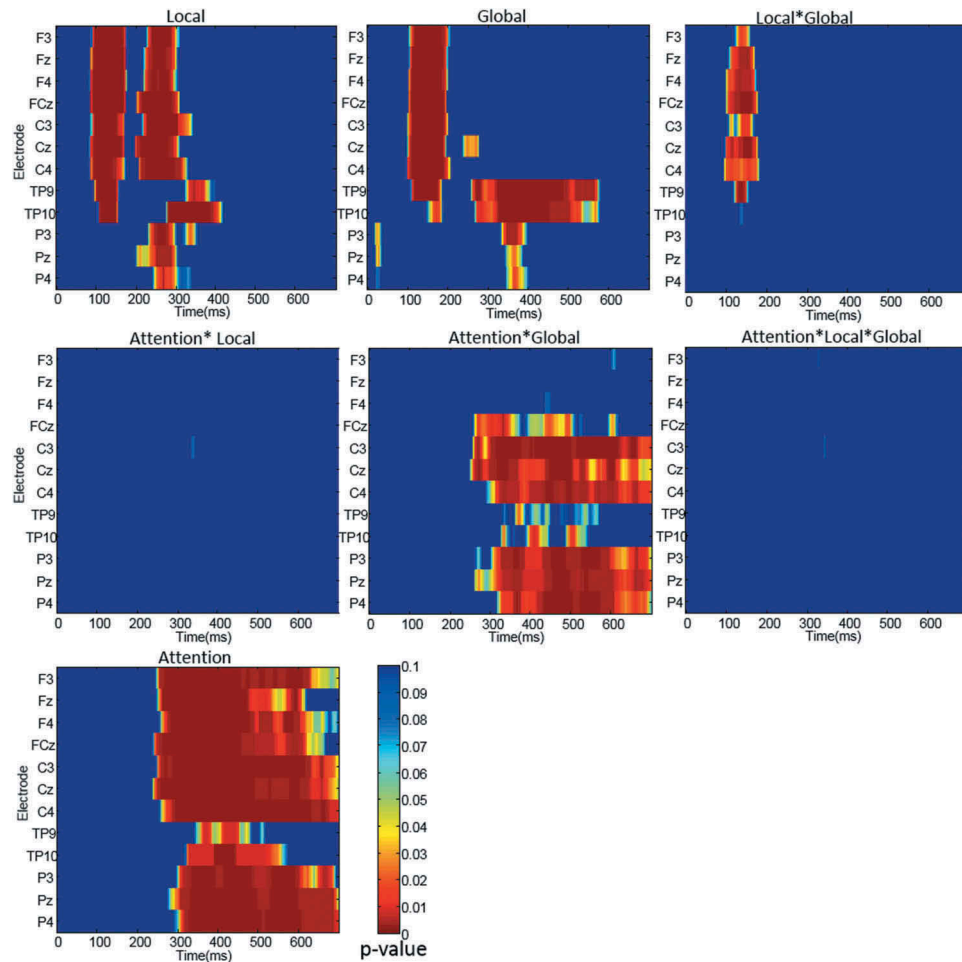
Figure 3. The timepoint-by-timepoint permutation analysis, showing p-values (family-wise error corrected) for the effects of the factors local, global, attention and their interactions. No timepoints were significant for Attention*Local and Attention*Local*Global analysis.

local–global studies that the global rule requires attention to be formed, they support the model of MMN as a 'modular' system which is isolated from higher-order attentional mechanisms (Ritter, Sussman, Deacon, Cowan, & Vaughan, 1999). Thus, while attention may reorganize the grouping of the patterns in the auditory stream (Sussman, 2007), the ability to represent longer patterns is not dependent on attention to the auditory modality.

As has been observed in earlier studies (Chennu et al., 2013; Bekinschtein et al., 2009; Strauss et al., 2015; Wacongne et al., 2011), the global-rule violation elicited a P3b wave only in the attended condition, confirming that the attention manipulation was successful. However, despite the absence of P3b for unattended global deviants, there was a significant interaction between the local and the global rule in the MMN timeframe which did not interact with the attention factor, indicating that even though the global-rule violations were not registered by the P3b-generating system, they influence the MMN. The results demonstrate that the processes indexed by the

negativity in the MMN timeframe are sensitive to both the local as well as the global status of a sound, and agree with other findings in showing the independence of attention and prediction in the N2 time range (see, e.g., Garrido, Rowe, Halász, & Mattingley, 2017).

While MMN is often referred to as attention-independent, relying on the observation that a robust MMN is elicited when the subjects are attending visual stimuli (Näätänen, Paavilainen, Rinne, & Alho, 2007), it has been suggested that attention can reorganize the auditory stream before it is passed on to the processes which lead to MMN generation (Sussman, Ritter, & Vaughan, 1998). Findings by Ritter et al. (1999) suggested that the system generating MMN was isolated from system underlying higher-level attentional processes which generate P3. The current results are in agreement with this model in the MMN time window, as the interaction between Local and Global factors did not show statistically significant three-way interaction with the Attention factor and we find that the empirical effect size for this interaction was negligibly

small ($\eta^2 < .001$). This further suggests that the formation of the global rule does not require conscious attention, in contrast to claims by earlier studies (Chennu et al., 2013; Bekinschtein et al., 2009; Strauss et al., 2015), but consistent with other designs exploring nested regularities (Fitzgerald & Todd, 2018; Todd et al., 2014). Importantly, the participants in the unattended condition were required to allocate their attentional resources to a demanding concurrent task on which they performed at a high level of accuracy. It is therefore highly unlikely that they were able to concurrently monitor and consciously detect concurrent longer-term patterns in the sequences. This indicates that the representation of the auditory environment in the order of several seconds and consisting of more complex patterns is carried out by the brain without needing the conscious attentional resources. The pattern of significant main effects of local and global as well as interaction between local and global factors seen at the frontocentral electrode was replicated in the mastoids, confirming that the observed pattern was not due to contamination from any other attentional effects which are typically expressed in the frontocentral, but not in the mastoid locations (Kujala, Tervaniemi, & Schröger, 2007).

The parallel representation of the local and global rule within the early, attention-independent detection system is consistent with the studies on multi-feature MMN, which have looked at the effect of violating predictions regarding different physical features of a sound independently (simple deviation) or in conjunction (abstract deviation). As demonstrated by a study (Takegata et al., 1999), representation of the simple and abstract rules was carried out in parallel, and the simultaneous deviation from both of these generated an MMN waveform which was explained by a model of the combined effect of the two types of rules. In the present case, the effect was interactive rather than additive, but the parallel representation of two rules in predicting the same feature is in agreement with this model. Here we show that the higher-order rules may consist of auditory gestalt representations with relatively long time-windows, as the representation of the global rule here needed to maintain at least the previous 5300 ms in order to hold two repetitions of the standard sequences (which may be considered to be the minimum standard-building requirement, see Cowan, Winkler, Teder, & Näätänen, 1993; Näätänen et al., 2007) and to react to the deviation in the final tone of the subsequent sequence.

An alternative interpretation of the results is that they reflect differences in the transitional probability of the tones between different conditions. The local deviant consisted of a physically different tone relative to the four preceding tones (XXXXY sequence). The interaction effect indicates that the ERP elicited by the final Y-tone had a different amplitude in the XXXXY block, where Y-tones made up 16% of all the tones, compared to the XXXXX block, where Y-tones made up 4% of all the tones. Thus, the difference in the amplitude would be consistent with the observation that the standard *vs.* deviant difference increases with the reduced overall probability of the deviant tone (Näätänen et al., 2007). By this interpretation, the effects would not be due to a global-rule interacting with a local rule but could represent factors such as stronger stimulus-specific adaptation in the XXXXX blocks. However, such an explanation does not agree with the significant effect of Global status on the negativity in the MMN timeframe in the XXXXX sequences. This effect was clearly demonstrated in our results by comparing the local standard-global deviant and local standard-global standard items. While the x-tone was locally standard in both of the blocks, the N2 amplitude to the final x-tone showed a significant difference in the amplitude depending on whether the sequence was embedded in an XXXXX-block or XXXXY-block, an effect also seen when examining the difference waveforms isolating the MMN component (Figure 1). Thus, the modulation of the negativity by the sequence-final X-tone could not be due to the local transitional probability relative to the previous tone but reflected the violation of the global rule representing the probability of the XXXXX-sequence relative to the block. This finding agrees with the literature demonstrating complex and longtime-window regularity representation in the auditory cortex, as indexed by MMN (Horvath et al., 2001; Paavilainen, 2013; Todd et al., 2014).

In contrast to the N2, the results in the P3a and P3b showed that their representation of violation of predictions was modified by attention. P3b was not sensitive to the local status of the sound. This has in earlier literature led to suggestion that P3b is uniquely representing the higher, block-level status of incoming information, which is dependent on attention, and is unable to operate when the attentional resources are removed (Chennu et al., 2013; Bekinschtein et al., 2009; Strauss et al., 2015). However, as discussed above, the modulation of the N2 by the global rule was clearly present in the attended as well as unattended condition. This indicates that the processes leading to negativity in the MMN timeframe could track the global status of the sound even under conditions where the P3b was not capable of representing it. Therefore, the data does not support the view that global-rule representation is uniquely performed by a system where violation generates P3b. Instead, the P3b appears to index

conscious, attentional processing of the attended sequences. This is consistent with the theories interpreting P3b as an index of detecting events which are salient or important relative to the currently maintained goal state (Polich, 2007; Walentowska, Moors, Paul, & Pourtois, 2016). This could also be the explanation for why the P3b was elicited only for the global deviants, but not for each local deviant in the blocks: the attentional grouping (Sussman, 2007) prioritized the five-tone sequences as salient or relevant.

The P3a, unlike P3b, was sensitive to the local status of the tone, similarly to N2, however we note that the effect size for the P3a ($\eta^2 = .02$) was smaller than the effect size of N2 ($\eta^2 = .21$), and the significance of the effect (p = .02) would not have survived correction for multiple comparisons. The finding of modulation by the local deviant, while small, is in agreement with findings of P3a following the violations which elicit the MMN, associated to orienting attention toward the deviant sound, or evaluating the contextual novelty of sounds (Escera & Corral, 2007). However, the effect was not modulated by the global rule, corroborating the suggestions that P3a generation is not directly dependent on the amplitude of MMN, instead it depends on processes happening prior to or parallel to MMN generation. Also, there was no significant interaction between Local and Attention factors: both attended and unattended local deviants elicited P3a with similar amplitude. However, similarly to P3b, there was an interaction between global status and attention: the effect of the global status on the amplitude change during deviant compared to standard was larger under the attention condition than visual distractor condition. As indicated by the time-point-by-timepoint permutation analysis, this may be due to two effects overlapping in the same time-window in the FCz electrode: an earlier Local effect which starts at ~200 ms post-deviant onset and extends from frontal to central electrodes and a later Attention*Global effect which begins ~250–300 ms post-deviant onset and is confined to frontocentral and central locations. Considering the distribution, it seems plausible that the early, post-200 ms component of the P3a, which is attention-independent and with more frontal location (Escera & Corral, 2007) is distinct from the later-onset component which appears to behave more similarly to P3b.

Future studies are needed to examine how complex rules can be learned by the MMN-generating representation system, and whether time-window length or other complexity measures of the pattern are central in determining the representation limits.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, 106(5), 1672–1677.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., … Bekinschtein, T. A. (2013). Expectation and attention in hierarchical auditory prediction. *Journal of Neuroscience*, 33(27), 11194–11205.

Cornella, M., Leung, S., Grimm, S., & Escera, C. (2012). Detection of simple and pattern regularity violations occurs at different levels of the auditory hierarchy. *PLoS One*, 7(8), e43604.

Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 909–921.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., … Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, 120(11), 1883–1908.

Escera, C., & Corral, M. (2007). Role of mismatch negativity and novelty-P3 in involuntary auditory attention. *Journal of Psychophysiology*, 21(3–4), 251–264.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160.

Fitzgerald, K., & Todd, J. (2018). Hierarchical timescales of statistical learning revealed by mismatch negativity to auditory pattern deviations. *Neuropsychologia*, 120, 25–34.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815–836.

Garrido, M. I., Rowe, E. G., Halász, V., & Mattingley, J. B. (2017). Bayesian mapping reveals that attention boosts neural responses to predicted and unpredicted stimuli. *Cerebral Cortex*, 28(5), 1771–1782.

Herholz, S. C., Lappe, C., & Pantev, C. (2009). Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience*, 10(1), 42.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, 69(3), 407–422.

Horvath, J., Czigler, I., Sussman, E., & Winkler, I. (2001). Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. *Cognitive Brain Research*, 12(1), 131–144.

Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Computational Biology*, 4(11), e1000209.

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology*, 74(1), 1–19.

Marti, S., Thibault, L., & Dehaene, S. (2014). How does the extraction of local and global auditory regularities vary with context? *PLoS One*, *9*(9), e107227.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590.

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

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.

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

Strauss, M., Sitt, J. D., King, J.-R., Elbaz, M., Azizi, L., Buiatti, M., … Dehaene, S. (2015). Disruption of hierarchical predictive coding during sleep. *Proceedings of the National Academy of Sciences*, *112*(11), E1353–E1362.

Sussman, E., Ritter, W., & Vaughan, H. G., Jr. (1998). Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research*, *789*(1), 130–138.

Sussman, E. S. (2007). A new view on the MMN and attention debate. *Journal of Psychophysiology*, *21*(3–4), 164–175.

Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neuroscience Letters*, *266*(2), 109–112.

Todd, J., Heathcote, A., Mullens, D., Whitson, L. R., Provost, A., & Winkler, I. (2014). What controls gain in gain control? Mismatch negativity (MMN), priors and system biases. *Brain Topography*, *27*(4), 578–589.

Todd, J., & Robinson, J. (2010). The use of conditional inference to reduce prediction error—A mismatch negativity (MMN) study. *Neuropsychologia*, *48*(10), 3009–3018.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, *108*(51), 20754–20759.

Walentowska, W., Moors, A., Paul, K., & Pourtois, G. (2016). Goal relevance influences performance monitoring at the level of the FRN and P3 components. *Psychophysiology*, *53*(7), 1020–1033.

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.