



Auditory Pattern Representations Under Conditions of Uncertainty—An ERP Study

Maria Bader*, Erich Schröger and Sabine Grimm

Cognitive and Biological Psychology, Institute of Psychology—Wilhelm Wundt, Faculty of Life Sciences, Leipzig University, Leipzig, Germany

OPEN ACCESS

Edited by:

István Winkler,
Hungarian Academy of Sciences
(MTA), Hungary

Reviewed by:

Anne Caclin,
Institut National de la Santé et de la
Recherche Médicale (INSERM),
France
Dean F. Salisbury,
University of Pittsburgh, United States
Petri Paavilainen,
University of Helsinki, Finland

*Correspondence:

Maria Bader
maria.bader@uni-leipzig.de

Specialty section:

This article was submitted to
Cognitive Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

Received: 19 March 2021

Accepted: 11 June 2021

Published: 09 July 2021

Citation:

Bader M, Schröger E and
Grimm S (2021) Auditory Pattern
Representations Under Conditions
of Uncertainty—An ERP Study.
Front. Hum. Neurosci. 15:682820.
doi: 10.3389/fnhum.2021.682820

The auditory system is able to recognize auditory objects and is thought to form predictive models of them even though the acoustic information arriving at our ears is often imperfect, intermixed, or distorted. We investigated implicit regularity extraction for acoustically intact versus disrupted six-tone sound patterns *via* event-related potentials (ERPs). In an exact-repetition condition, identical patterns were repeated; in two distorted-repetition conditions, one randomly chosen segment in each sound pattern was replaced either by white noise or by a wrong pitch. In a roving-standard paradigm, sound patterns were repeated 1–12 times (standards) in a row before a new pattern (deviant) occurred. The participants were not informed about the roving rule and had to detect rarely occurring loudness changes. Behavioral detectability of pattern changes was assessed in a subsequent behavioral task. Pattern changes (standard vs. deviant) elicited mismatch negativity (MMN) and P3a, and were behaviorally detected above the chance level in all conditions, suggesting that the auditory system extracts regularities despite distortions in the acoustic input. However, MMN and P3a amplitude were decreased by distortions. At the level of MMN, both types of distortions caused similar impairments, suggesting that auditory regularity extraction is largely determined by the stimulus statistics of matching information. At the level of P3a, wrong-pitch distortions caused larger decreases than white-noise distortions. Wrong-pitch distortions likely prevented the engagement of restoration mechanisms and the segregation of disrupted from true pattern segments, causing stronger informational interference with the relevant pattern information.

Keywords: auditory processing, P3a, complex sound patterns, event-related potentials, mismatch negativity

INTRODUCTION

Acoustic information, which arrives at our ears and informs us about objects in the outer world, is often imperfect. Parts of the relevant object information might be obscured by extraneous noise from concurrent auditory sources, for instance, from some sudden, interfering background sounds like a honking car or a barking dog. In other instances, parts of the relevant auditory objects can be missing. Furthermore, particularly in the domains of speech and music, mistakes in production might lead to imperfect recurrences of the same object, for instance, when a familiar melody is played with a missing or a wrong note or a word is uttered with an incorrect phoneme. Such

occurrences can intermix with or distort the acoustic input we are currently paying attention to, and might impede its intelligibility.

Our ability to deal with such distortions and to interpret the acoustic environment, even if it is degraded to a certain degree, is a remarkable asset. One should note that, on the one hand, our auditory system is able to detect even slight variations in the acoustic input. Previous studies found that the auditory system is sensitive to very subtle acoustic changes (e.g., in sound frequency), particularly if the respective varying sound event occurs after a row of invariant repetitions (Sams et al., 1985; Winkler et al., 1990). On the other hand, our brain must be able to tolerate variation to some extent. That is, we need to neglect input variation that is irrelevant for the current task or we need to compensate for missing or distorted information. Usually, we are well able to maintain a stable representation of objects in the environment, even if we encounter them occasionally in a degraded form. In fact, mechanisms of perceptual prediction and restoration help us to fill in and reconstruct occluded or obscured information. This occurs not only in the visual world (e.g., in the case of the blind spot; Walls, 1954; Ramachandran, 1992; De Weerd, 2006; Spillmann et al., 2006) but also in the auditory domain. Studies on the continuity illusion show that a tone or a word containing a short gap may be perceived as continuous if the gap is filled with noise. The listener then perceives the missing information, suggesting that the auditory system predicts and interpolates through the absent information (Warren et al., 1997; Micheyl et al., 2003; Riecke et al., 2007; Shahin et al., 2009; Bendixen et al., 2014).

Especially in speech and music research, perceptual restoration and filling-in processes are attributed to top-down influences (DeWitt and Samuel, 1990; Shinn-Cunningham and Wang, 2008). When phonemes are replaced with white noise, a word utterance can still seem intact, and, oftentimes, listeners cannot say which part of the uttered sentence was missing (Warren and Warren, 1970). Although this effect is clearly influenced by knowledge and experience, part of the phenomena might occur pre-attentively on lower levels of processing (Micheyl et al., 2003; Riecke et al., 2007).

Electrophysiological markers like the mismatch negativity (MMN) can serve as an index for such implicit and automatic compensatory mechanisms of early processing stages (Micheyl et al., 2003). The MMN is elicited by sounds violating a detected regularity in a sequence of sounds, for example, when a tone differing in pitch (deviant) is presented following a sequence of tones with identical pitch (standards). MMN reflects the process of deviance detection, and its presence indirectly implies that the regularity inherent in the standard tones has been encoded (Winkler, 2007). Although early accounts interpreted the MMN as the outcome of a retrospective comparison process between regularity representations and the incoming deviant sound, newer accounts emphasize that regularity representations are part of an internal model, prospectively generating predictions about future input (Wacongne et al., 2012; Bendixen et al., 2014; Winkler and Schröger, 2015). Previous studies showed that predictions and regularity representations of standards build up even in the absence of exact repetitions of a stimulus, for instance, when using abstract regularities (Tervaniemi et al., 2000;

Brattico et al., 2006; Bendixen and Schröger, 2008; Bader et al., 2017) or in the presence of noise that degrades the physical information (Muller-Gass et al., 2001; Micheyl et al., 2003; Kozou et al., 2005). All these studies demonstrate the tolerance of the MMN systems to a considerable amount of variability in the sequence of standard sounds, including cases of imperfect repetitions of standard sounds.

These findings also align with studies on auditory object segregation. For example, McDermott et al. (2011) showed that the auditory system quickly recognizes invariant patterns, even when they are embedded in a changing acoustic background of competing sounds. The authors suggest a mechanism of cross-correlating dynamic spectrotemporal input patterns, which filters for invariances between different occurrences of the same auditory event (despite its being mixed with background noise). Nevertheless, pattern recognition is impaired for sound patterns that are not identically repeated within the mixtures of changing backgrounds. In event-related potentials (ERP) studies, deviant sounds elicit an MMN of decreased amplitude and increased latency in cases where regularity formation is impeded by abstract variations or noise masking (Muller-Gass et al., 2001; Niemitalo-Haapola et al., 2015; Bader et al., 2017).

The MMN component can be followed by a P3a component distributed over fronto-central scalp regions. Typically, P3a is interpreted as signaling an involuntary attentional switch from a primary task to the deviant stimulus (Squires et al., 1975; Escera et al., 2000; Friedman et al., 2001; Wetzel and Schröger, 2007). P3a is sometimes also discussed to reflect a higher level but automatic evaluation of novelty rather than the switch of attention itself (Horvath et al., 2008; Wetzel et al., 2013; Winkler and Schröger, 2015). The P3a amplitude is modulated by cognitive and working memory demands of the task (Berti et al., 2004). P3a amplitude and latency can be modulated by experimental manipulations affecting salience, such as in the presence of abstract variations (Bader et al., 2017) and quality degradations of the regularity inherent in the standards, for example, by added noise, interruptions, or frequency distortions (Micheyl et al., 2003; Bader et al., 2017; Uhrig et al., 2017). That is, P3a amplitude might be decreased and latency increased if deviants violate a regularity that is not defined by exact sound or pattern repetitions, for instance, in the case of a new melodic pattern occasionally occurring among pattern repetitions in a transposed form (Bader et al., 2017).

In the current study, we investigated the early phase of implicit auditory pattern learning. The presented sounds were attended; however, the pattern regularity rule was irrelevant for the task at hand. Thus, learning occurred incidentally rather than intentionally (Perruchet and Pacton, 2006). During the experiment, a standard six-tone pattern is either repeated identically to the listeners or repeated, containing distortions, which leave only 5/6 of the standard pattern intact during each presentation. In particular, the three experimental conditions included (1) an exact-repetition condition (exact) in which sound patterns within a train were perfect repetitions without any distortion; (2) a white-noise-distortion condition (wn) in which one tonal segment of the repeating sound pattern was replaced randomly by Gaussian noise, and (3) a wrong-pitch-distortion

condition (wp) in which one tonal segment of the repeating sound pattern underwent a random shift in pitch. The respective distortions could affect a different tonal segment in each pattern presentation; the same segment would be intact in the next pattern presentation. Importantly, we were not measuring the ERPs to the distortions in every trial but, rather, the ERPs to occurrences of a completely new pattern, whether or not the preceding pattern learning happened with or without distortions. If the buildup of a pattern representation is (at least partly) robust against pattern variance in complex auditory stimuli, we hypothesize that MMN and P3a will be elicited in conditions with and without distortions.

However, given that a lower quantity of unambiguous information guides the initial learning process, one might expect poorer pattern representation from distorted standard sounds as indicated by decreased MMN and P3a amplitude and possible latency delays. Given that an equal portion (five out of six tones) was unambiguously informative about the identity of the standard pattern in both distortion conditions, one might expect to find similar impairments in the two distortion conditions. Nevertheless, there are certainly qualitative differences between white noise and wrong pitch distortions. Noise contains the frequencies of the “true” pattern segment and can involve filling-in or restoration processes, once an internal representation of the standard pattern has been formed; whereas wrong pitch distortions might interfere with the formation of a pattern representation. Therefore, differential effects on MMN and/or P3a could be expected in the two distortion conditions, particularly with stronger impairments in the wrong-pitch condition. If that is the case, one can conclude that potentially different mechanisms rule the formation of pattern identity representations in the context of white noise or wrong pitch distortions. This would also be compatible with masking studies, showing stronger degrading effects on speech intelligibility when it is masked by similar stimulus material, such as irrelevant speech (as in the case of informational masking), than when it is masked by noise (as in the case of energetic masking) (Cooke et al., 2008; Lidestam et al., 2014).

Furthermore, if we observed such distortion-specific effects already at the level of MMN, this would argue in favor of an impact of our manipulation on relatively early merely perceptual and automatic processing levels. In contrast, if they were confined to the P3a level, this would indicate that distortion processing occurs only at a later stage in the course of more context-dependent novelty evaluation and automatic orienting of attention.

As mentioned above, we expect MMN and P3a elicitation to full pattern changes in all conditions and a possible differentiable effect of the inserted distortions. These effects might also interact with the number of preceding standards. Therefore, the number of standard patterns (with or without distortions) preceding a deviant was varied (1, 2, 3, 6, or 12 standard patterns) in a roving-standard paradigm (Winkler et al., 1996; Bendixen and Schröger, 2008; Garrido et al., 2009). According to previous studies (Cowan et al., 1993; Winkler et al., 1996; Bendixen et al., 2007; Garrido et al., 2008; Bader et al., 2017), we expect to see a growth in MMN (and P3a) amplitude with increasing numbers of preceding

standards in all conditions, if the implicit learning mechanisms are robust against the inserted distortions. Likely, both standard and deviant ERPs contribute to the MMN increase with deviant ERPs growing more negative and standard ERPs growing more positive (and *vice versa* for the P3a) as a function of the number of previous standards in a train (Baldeweg et al., 2004; Bendixen et al., 2007; Costa-Faidella et al., 2011; Bader et al., 2017). We will model the emergence and growth, using a logarithmic regression analysis, because regression coefficients can be informative of the time course and the strength of the implicit formation of pattern representations in our three distortion conditions.

MATERIALS AND METHODS

Participants

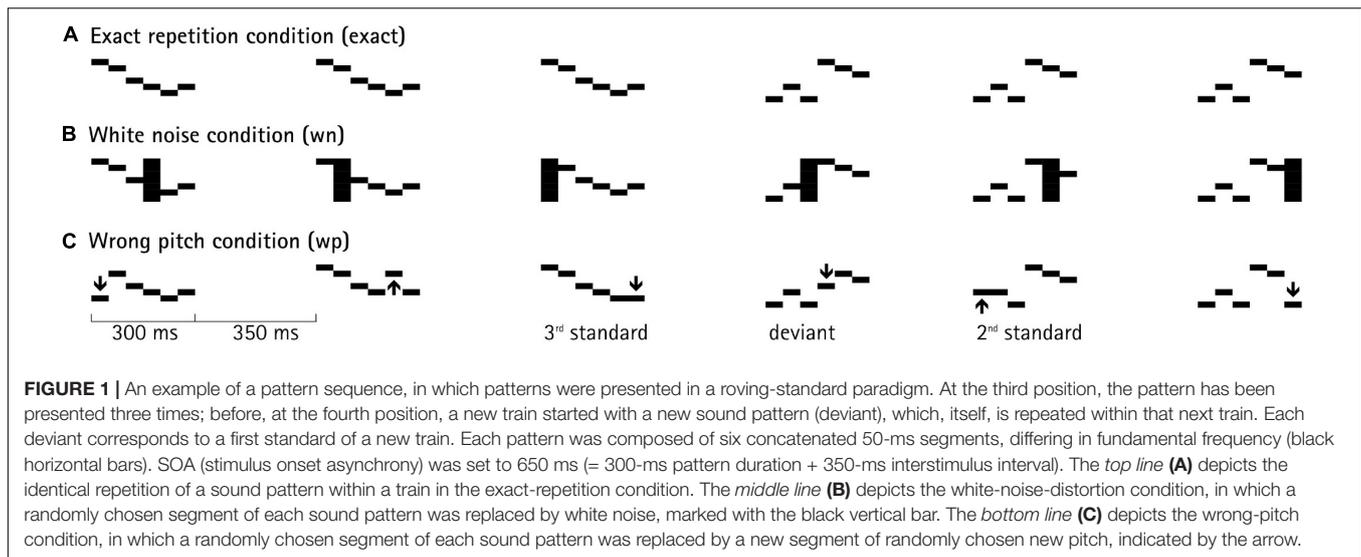
The experimental protocol was approved by the Ethical Committee of the University of Leipzig and was in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The participants gave written informed consent before experimental sessions. All the subjects in the experiment participated for credit points or monetary compensation (€8 per hour). Of the 19 healthy subjects (age range: 19–40 years, 18 females) that participated in this research, all reported normal hearing and 17 out of 19 participants were right-handed. All of them were Leipzig University students, and 84% reported to have played a musical instrument for some time ($M = 6.6$ years, $SD = 4.2$ years).

Materials

As in one of our previous studies (Bader et al., 2017), auditory stimuli were composed of 300-ms sound patterns, consisting each of six concatenated 50-ms segments with randomly chosen fundamental frequencies between 220 and 880 Hz (in 25 semitone steps). Harmonics were added to each fundamental frequency until a cutoff at 6,000 Hz. Starting at 3,000 Hz, tonal segments were modulated by reducing the signal linearly resulting in 0% intensity at 6,000 Hz. For a smoother sound, odd harmonics (uneven positive integer multiples of the fundamental frequency) were additionally attenuated to 20% of their intensity. To prevent loudness differences between segments, intensities were root mean square equalized. Segments included a 5-ms rise and a 5-ms fall time, and there were no gaps introduced between the six segments when concatenating them to a sound pattern. The stimulus onset asynchrony (SOA) between patterns presented in our auditory sequences was set to 650 ms.

Design and Procedure

Sound patterns were presented in a roving-standard paradigm (Cowan et al., 1993; Winkler et al., 1996; Bendixen et al., 2007; Garrido et al., 2008) with varying train lengths; a randomly generated sound pattern was presented either 1, 2, 3, 6, or 12 times in a sequence before a newly generated pattern occurred, which started a new train of stimuli. In a single block of the experiment, each possible train length occurred 10 times in random order, resulting in 240 pattern presentations per block. As in our previous study (Bader et al., 2017), 10 additional trials



of train length 1 were included in each block, in the way that one trial of train length 1 was directly followed by another trial of train length 1. In this context, three pattern changes always occurred in a row. This served to have pattern changes that did not follow a pattern repetition, and it ensured the investigation of memory trace formation, starting with a first pattern presentation. Here, the first pattern change served as a deviant with respect to the previous train, the second served as a “standard” of train length 1, and the third served as a “deviant” of train length 1. This terminology is consistent with the one used for other train lengths but, admittedly, arbitrary, since stimuli of train length 1 do not have an actual history of pattern repetition. Overall, for each train length, a similar number of standard and deviant patterns was available for ERP analysis (Bader et al., 2017).

The experiment consisted of two sessions with a total of 36 blocks (12 blocks \times 3 conditions). In each session, six blocks of each condition were presented in random order. In the exact-repetition condition, patterns were repeated 100% identically within a train. In the white-noise condition, during each presentation of a melodic pattern, one randomly chosen segment out of the six was replaced by a 50-ms snippet of white noise. In the wrong-pitch condition, the pitch of one segment in each pattern was changed to a randomly chosen new pitch, while keeping all other characteristics of that segment (e.g., timbre) unchanged. The randomly chosen pitch for the new segment could keep the contour of the pattern intact, or violate it with equal probability. Each position of a pattern (1–6) could be affected by this manipulation with equal probability in the two distortion conditions—randomly selected from trial to trial. The intensity of the white noise segment was root-mean-square equalized to the rest of the pattern. **Figure 1** depicts an example of a train with three patterns and the beginning of the following train.

At the beginning of the first session, ability of the participants to tell whether two melodic phrases are the same or different was measured *via* the melody part of musical ear test (MET) of Wallentin et al. (2010). It contains 52 trials during which two

short melodies (comprising of three to eight tones) are played with a tempo of 100 beats per minute, one after the other, with sampled piano sounds. This part of the MET lasts approximately 10 min. In half of the 52 trials, the two melodies are identical. In order to be above-chance level, the participants must score 32 out of 52 trials correctly (= 62%).

During the experimental sessions, the participants were seated in an electrically and acoustically shielded chamber in our laboratory at the Institute of Psychology of Leipzig University. To minimize eye movements, the participants fixated a cross on a computer screen placed behind a window outside the chamber 130 cm from the eyes of the participants. Auditory stimuli were presented binaurally over headphones (Sennheiser HD 25) at an intensity level of approximately 78 dB SPL. The participants were not informed about the roving rule. While listening to the presented sound patterns, the subjects performed a loudness change detection task in order to ensure their attention on the auditory stimulation. Occasionally, the sound patterns were presented with a higher volume (+ 4 dB, five sound patterns = 2% per block) or with a lower volume (–4 dB, five sound patterns = 2% per block). The participants pressed the left button of a response pad as soon as they detected a sound pattern of lower volume and the right button of the response pad as soon as they detected a sound pattern of higher volume. The targets were distributed randomly over each block with the restriction of at least two non-targets in between. After finishing a block, the participants received feedback on their performance (a ratio of hits, interchanged buttons, false alarms, and their mean reaction time). After each block, the subjects had a short break, allowing for movements.

Additionally, at the end of the second experimental session, the participants performed an active pattern change detection task to measure behavioral detectability of pattern changes. In these active blocks, the SOA was prolonged to 1,100 ms for the participants to have sufficient time for solving the task. They were instructed to detect the onset of a new train while ignoring segment distortions. They pressed a button of the response pad as

soon as they detected a change to a new sound pattern. After short training, each subject performed one block of each distortion condition, consisting of 200 trials each (8 times train lengths 2, 3, 6, and 12; 16 times train length 1). The order of conditions was counterbalanced over the participants.

Data Acquisition and Analyses

Electroencephalography data were collected continuously from 64 Ag/AgCl active electrodes. The electrodes were positioned according to the international 10–20 system in a nylon cap. The vertical and horizontal electrooculograms (EOG) were measured with external electrodes placed above and below the right eye and at the outer canthi of both eyes, respectively. As possible offline references, additional electrodes were placed on the tip of the nose and over each mastoid. All electrode signals were (DC) amplified and continuously sampled with a rate of 512 Hz (an anti-aliasing filter with -3 dB at $1/4$ of the sampling frequency) by BioSemi Active-Two amplifiers. No high-pass filter was applied online. The BioSemi system uses the common mode sense (CMS) and the driven right leg (DRL) electrode—placed at different sites at the back of the head—to reference the recording to the CMS–DRL ground while minimizing the effect of external noise sources¹.

Offline, EEG data were re-referenced to the average signal of the two mastoids (Paavilainen et al., 1989; Ritter et al., 1992; Schröger, 1997) and filtered using a 0.5 Hz high pass filter (transition bandwidth of 1 Hz, filter order 1,690) and a 35 Hz low pass filter (transition bandwidth of 10 Hz, filter order 170). Both filters were Kaiser windowed sinc FIR filters (beta = 5.653, stopband attenuation = -60 dB implemented in EEGLAB (Delorme and Makeig, 2004; Widmann et al., 2015). Epochs of 650-ms duration were extracted from the continuous electroencephalography (EEG) time-locked to sound pattern onset. No baseline correction was applied to avoid the introduction of pre-stimulus neuroelectric activity in the baseline period into the post-stimulus waveforms (Urbach and Kutas, 2006). Epochs were sorted for each participant, condition, and stimulus type. Artifactual epochs with a signal range exceeding $100 \mu\text{V}$ on any recording channel (including EOG channels) were discarded from the analyses. To exclude trials containing artifacts, which are not characterized by extreme amplitude but by noise, we additionally ran a sorted averaging procedure (Muhler and von Specht, 1999; Rahne et al., 2008), during which all epochs of one condition and one participant were sorted according to their noise level (quantified by the root mean square of the voltage and sorted from lowest to highest) and successively entered the average as long as they increase the overall signal-to-noise ratio of the average. Epochs with extreme noise levels, which have a deteriorating effect on the signal to the noise ratio, were excluded. On average, 87 epochs (= 73%) remained for analysis across conditions and the participants. The data of all the participants went into analyses.

Event-related potentials were averaged for deviant sounds (the first pattern of a new train) and for standard sounds (the last corresponding pattern of the preceding train) in each

condition. Difference waves were computed by subtracting the standard from the deviant ERPs. To compare ERP effects, in general, between the three conditions, we pooled the deviant and standard ERPs of train lengths 3, 6, and 12 in order to extract a general difference waveform in each condition (general analysis). Train lengths 1 and 2 were not included in the general analysis, because former studies showed that perceptual regularity extraction requires at least two repetitions (i.e., three presentations) of the standard before unexpected complex sound patterns elicit an MMN (Bendixen and Schröger, 2008; Bader et al., 2017). With non-parametric cluster-based permutation tests (Maris and Oostenveld, 2007), we tested for the presence of significant MMN and P3a components in each condition within this general analysis by comparing the whole epoch of the deviant ERP with the standard ERP (α -level for channels and clusters: Monte Carlo $p < 0.05$). About 1,000 permutations were run for each test, and dependent samples t -tests quantified the effect. These analyses and the creation of topographical scalp plots were run in FieldTrip (Oostenveld et al., 2011).

For the statistical assessment, amplitude measures were derived from the individual ERPs as the mean signal amplitude in the component time interval for combined conditions (MMN: 200–300 ms after the stimulus onset, P3a: 400–500 ms after the stimulus onset). An estimate of relative slope of each component at the latency for which 75% of the peak amplitude was reached was determined, using the jackknife approach (Miller et al., 1998; Kiesel et al., 2008).

The MMN and P3a amplitude were separately subjected to a repeated measures analysis of variance (ANOVA) with the factors condition (exact, wn, wp), stimulus type (deviant, standard), and train length (1, 2, 3, 6, and 12) as within-subject factors. For the MMN, maximal peak deflections were distributed over different electrodes in central-parietal regions (CPz, Pz, POz). Following a data-driven approach, we focused our MMN analyses on this region with maximal amplitude, and we discuss this rather unexpected topography later (see Discussion). To examine possible topography differences, the factor electrode (CPz, Pz, and POz) was added to the ANOVA. P3a measures were taken from electrode FCz because amplitude deflected maximally at this electrode in all conditions. Three condition-wise repeated measures ANOVAs, with the factors stimulus type (deviant, standard) and train length (1, 2, 3, 6, and 12), and two stimulus-type-wise repeated measures ANOVAs, with the factors condition (exact, wn, and wp) and train length (1, 2, 3, 6, and 12), were run to explore the three-way interactions for MMN and P3a. To investigate in more detail whether and how amplitude of MMN and P3a, as well as the contribution of deviant and standard stimuli to amplitude of both components, changed systematically as a function of train length; *post hoc* logarithmic trend analyses were run for difference waveforms, deviant and standard ERPs, separately. The standardized regression coefficient r and its 95% confidence interval (lower and upper bounds) were reported. All ANOVAs generalized eta squared (η^2) served as an estimate of effect size (Bakeman, 2005), and the Greenhouse–Geisser correction was applied when the assumption of sphericity was violated (corrected d fs were reported). All parametric statistical analyses

¹<https://www.biosemi.com/faq/cms&drl.htm>

and the creation of figures were run with RStudio (RStudio. PBC. Version 1.3.1073).

To ensure the participants attended to the auditory stimulation during the EEG session, we analyzed sensitivity d' according to Macmillan and Creelman (2004) to the loudness changes. Sensitivity d' was adjusted for extreme hit or false alarm rates to avoid infinite d' values by adding 0.5 to all response counts (Brown and White, 2005; Hautus and Lee, 2006). Reaction times were measured by calculating the latency between the pattern onset and the key press. Response latencies greater than two SOAs (1,300 ms) were excluded from analysis. To check further for unwanted side effects, the behavioral data in the loudness change detection task were additionally analyzed in repeated measures ANOVA with the factors session (first, second) and condition (exact, wn, and wp).

The active pattern change detection task at the end of the second EEG session was analyzed in terms of the signal detection theory, by extracting an index of sensitivity d' (Macmillan and Creelman, 2004) and correcting it for an extreme hit or false alarm rates by adding 0.5 to all response counts (Brown and White, 2005; Hautus and Lee, 2006). Reaction times were measured by calculating the latency between the pattern onset and the key press. Responses with latencies greater than the SOA (1,100 ms) were excluded from the analysis. To compare the behavioral performance between conditions, repeated measures ANOVAs and Bonferroni-corrected two-tailed t -tests were run. Cohen's d was calculated as an estimate of effect size for Student's t -tests. Additionally, hit rates were analyzed *via* repeated measures ANOVA with the factors condition (exact, wn, and wp) and train length (1, 2, 3, 6, and 12) and *via* logarithmic trend analysis to investigate behavioral effects condition-wise and as a function of the number of previous standard repetitions.

RESULTS

Behavioral Performance in the MET and in the Loudness Change Detection Task

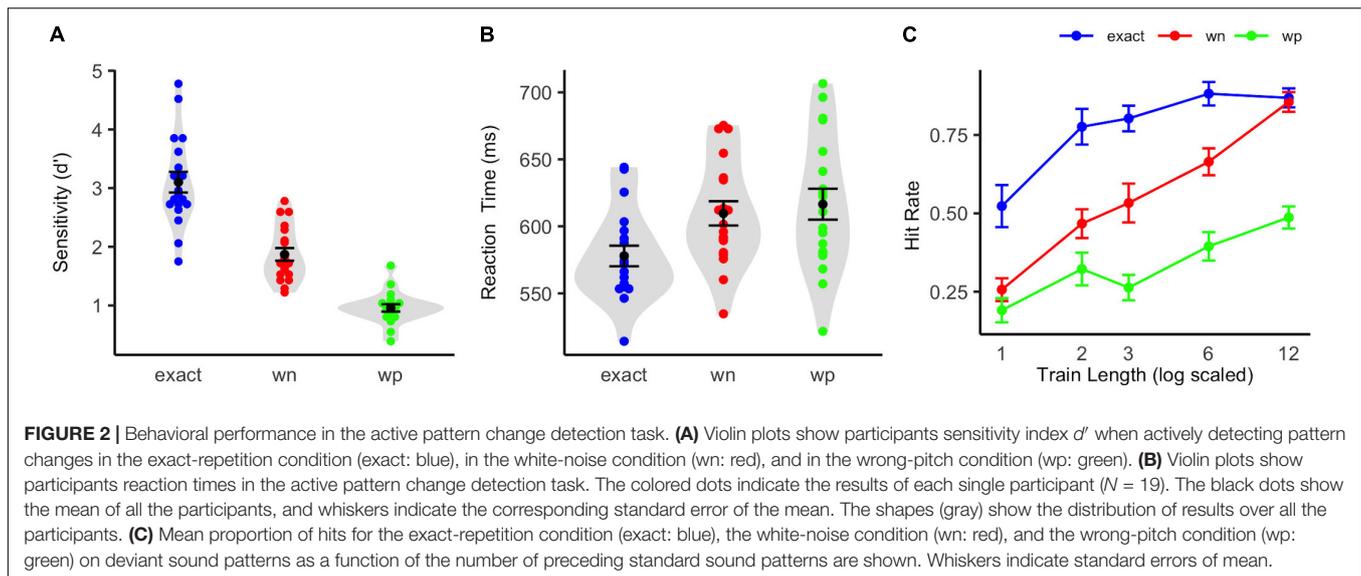
In the MET, the participants scored, on average, 72% correct (range: 42–88% $SD = 12\%$, 17/19 subjects scored above chance). In the loudness change detection task, which participants performed during the EEG recordings, the targets were discriminated with high accuracy. In the first session, averaged sensitivity over the participants ($N = 19$) was $d' = 3.71$ ($SD = 0.58$) in the exact-repetition condition, $d' = 3.45$ ($SD = 0.64$) in the white-noise condition, and $d' = 3.60$ ($SD = 0.70$) in the wrong-pitch condition. In the second session, averaged sensitivity over the participants ($N = 19$) was $d' = 3.92$ ($SD = 0.52$) in the exact-repetition condition, $d' = 3.80$ ($SD = 0.51$) in the white-noise condition, and $d' = 3.92$ ($SD = 0.58$) in the wrong-pitch condition. A repeated measures ANOVA with the factors session (first, second) and condition (exact, wn, and wp) revealed the main effect of session [$F(1, 18) = 39.17$, $p < 0.001$, $\eta^2 = 0.69$] but no significant effect of condition [$F(2, 36) = 3.07$, $p = 0.060$, $\eta^2 = 0.15$]. No interaction between condition and sessions was found [$F(2, 36) = 0.52$, $p < 0.60$, $\eta^2 = 0.03$].

In the first session, reaction times for correctly detected target sounds were $M = 634$ ms ($SD = 50$ ms) in the exact-repetition condition, $M = 664$ ms ($SD = 53$ ms) in the white-noise condition, and $M = 631$ ms ($SD = 49$ ms) in the wrong-pitch condition. In the second session, reaction times for correctly detected target sounds were $M = 609$ ms ($SD = 59$ ms) in the exact-repetition condition, $M = 628$ ms ($SD = 58$ ms) in the white-noise condition, and $M = 617$ ms ($SD = 54$ ms) in the wrong-pitch condition. A two-way repeated measures ANOVA revealed a significant effect of condition [$F(2, 36) = 13.99$, $p < 0.001$, $\eta^2 = 0.04$] and of session [$F(1, 18) = 14.08$, $p = 0.001$, $\eta^2 = 0.05$]. These main effects were not qualified by an additional interaction [$F(2, 36) = 1.32$, $p = 0.28$, $\eta^2 = 0.01$]. Bonferroni-corrected two-tailed t -tests showed that, across the sessions, the reaction times in the white-noise condition differed significantly from the exact-repetition condition and the wrong-pitch condition [exact vs. wn: $t(37) = -3.59$, $p = 0.003$, $d = -1.18$; wn vs. wp: $t(37) = 3.09$, $p = 0.001$, $d = 1.02$; exact vs. wp: $t(37) = -0.43$, $p = 1.000$, $d = -0.14$].

The participants showed a high sensitivity in the active pattern change detection task at the end of the EEG recordings in the exact-repetition condition [$d' = 3.12$ ($SD = 0.77$)]. The sensitivity in the white-noise condition was $d' = 1.90$ ($SD = 0.47$) and in the wrong-pitch condition $d' = 0.96$ ($SD = 0.28$). Repeated measures ANOVA comparing the three conditions revealed a significant effect of condition [$F(0.96, 34.53) = 125.68$, $p < 0.001$, $\eta^2 = 0.74$]. Bonferroni-corrected two-tailed t -tests showed significant differences between each condition pair [exact vs. wn: $t(18) = 8.47$, $p < 0.001$, $d = 3.99$; wn vs. wp: $t(18) = 10.12$, $p < 0.001$, $d = 4.77$; exact vs. wp: $t(18) = 13.31$, $p < 0.001$, $d = 6.27$].

On average, reaction times for correctly detected target sounds were $M = 578$ ms ($SD = 34$ ms) in the exact-repetition condition, $M = 610$ ms ($SD = 39$ ms) in the white-noise condition, and $M = 617$ ms ($SD = 50$ ms) in the wrong-pitch condition. A one-way repeated measures ANOVA revealed a significant effect of condition [$F(2, 36) = 12.54$, $p < 0.001$, $\eta^2 = 0.15$]. Bonferroni-corrected two-tailed t -tests revealed significant differences between exact vs. wn: $t(18) = -3.94$, $p = 0.003$, $d = -1.86$ and exact vs. wp: $t(18) = -3.90$, $p < 0.003$, $d = 1.84$. The difference in reaction times between the white noise and the wrong-pitch condition was not significant [$t(18) = -1.08$, $p = 0.88$, $d = 0.51$]. The results of sensitivity and reaction time analysis can be seen in **Figures 2A,B**.

A repeated measures ANOVA on the hit rates, including the factors condition (exact, wn, and wp) and train length (1, 2, 3, 6, and 12) showed a significant main effect of condition [$F(2, 36) = 103.98$, $p < 0.001$, $\eta^2 = 0.48$] and a significant main effect of train length [$F(4, 72) = 51.36$, $p < 0.001$, $\eta^2 = 0.35$]. These main effects were qualified by an additional interaction [$F(8, 144) = 5.11$, $p < 0.001$, $\eta^2 = 0.08$]. A highly significant logarithmic trend for the train length effect was revealed for each condition [exact: $F(1, 18) = 22.52$, $p < 0.001$, $\eta^2 = 0.56$; wn: $F(1, 18) = 228.30$, $p < 0.001$, $\eta^2 = 0.93$; wp: $F(1, 18) = 35.10$, $p < 0.001$, $\eta^2 = 0.66$]. Logarithmic regression analyses revealed a steeper increase of the hit rates with increasing train length in the white-noise condition than in the wrong-pitch condition, as



confidence intervals did not overlap (wn: $r = 0.23$; CI: 0.19–0.28, wp: $r = 0.11$; CI: 0.07–0.16). A ceiling effect in the exact-repetition condition caused a non-optimal fit of the logarithmic regression and, therefore, an overlap with the confidence interval of the wrong-pitch condition, but not so with the white-noise condition ($r = 0.13$; CI: 0.08–0.18). Mean hit rates are depicted in **Figure 2C**.

EEG Data

The non-parametric cluster-based permutation tests (see **Figure 3**) showed that grand-averaged difference waveforms (collapsed for train length 3, 6, and 12) elicited negative deflections prior to 300 ms after the stimulus onset in all conditions, likely reflecting MMN, even though its amplitudes were largest at central-parietal and parietal electrodes (exact: Pz: $M = -1.81 \mu\text{V}$ at 252 ms; wn: CPz: $M = -1.01 \mu\text{V}$ at 240 ms; wp: POz: $M = -0.91 \mu\text{V}$ at 211 ms). This negative component was followed by a positive deflection in all conditions, peaking at electrode FCz (exact: $M = 3.90 \mu\text{V}$ at 402 ms; wn: $M = 1.83 \mu\text{V}$ at 475 ms; wp: $M = 1.79 \mu\text{V}$ at 488 ms), reflecting P3a (see **Figure 4B**).

MMN Latency

Jackknife estimates of the MMN slope latency of the grand-averaged difference waves (at collapsed electrodes CPz, Pz, and POz), using the relative 75% peak amplitude criterion of the grand-averaged difference waves were $M_{jack} = 223$ ms ($SD_{jack} = 3.12$ ms) for the exact-repetition condition, $M_{jack} = 175$ ms ($SD_{jack} = 14.38$ ms) for the white-noise condition and $M_{jack} = 194$ ms ($SD_{jack} = 2.31$ ms) for the wrong-pitch condition. A one-way repeated measures ANOVA did not reveal a significant effect of condition: $F_{adj}(1.11, 19.93) = 0.44$, $p = 0.51$, $\eta^2 = 0.09$.

P3a Latency

Jackknife estimates of the P3a slope latency at electrode FCz of relative 75% peak amplitude criteria of the difference were

$M_{jack} = 369$ ms ($SD_{jack} = 2.46$ ms) for the exact-repetition condition, $M_{jack} = 402$ ms ($SD_{jack} = 1.19$ ms) for the white-noise condition and $M_{jack} = 428$ ms ($SD_{jack} = 7.08$ ms) for the wrong-pitch condition. A one-way repeated measures ANOVA did not reveal a significant condition effect: $F_{adj}(1.19, 21.45) = 2.68$, $p = 0.12$, $\eta^2 < 0.001$.

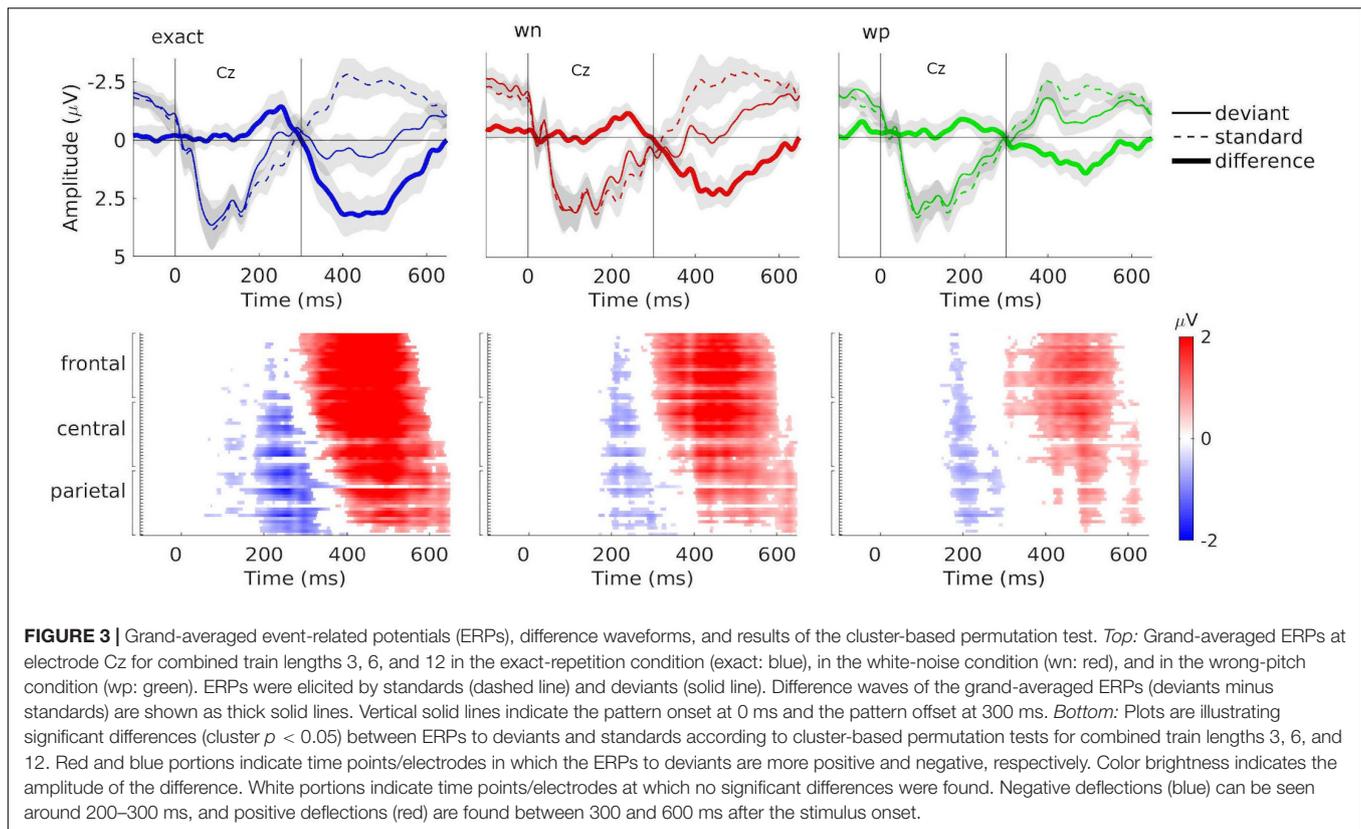
MMN Mean Amplitudes

ERP difference waves and scalp distributions of MMN (and P3a) can be seen in **Figure 4**.

A four-way repeated measures ANOVA with the factors condition (exact, wn, wp), stimulus type (dev, stand), train length (1, 2, 3, 6, and 12) and electrode (CPz, Pz, and POz) for mean amplitudes in the MMN time window (200 to 300 ms after the stimulus onset) neither revealed a significant 4-way interaction [$F(6.36, 114.44) = 1.50$, $p = 0.18$, $\eta^2 < 0.001$] nor any three-way interaction, including the factors electrode and condition [condition \times type \times electrode: $F(2.76, 49.64) = 1.05$, $p = 0.38$, $\eta^2 < 0.001$; condition \times train length \times electrode: $F(5.48, 98.59) = 1.12$, $p = 0.36$, $\eta^2 < 0.001$] that would point to topographical differences between conditions.

Instead, the ANOVA revealed significant main effects [condition: $F(2, 36) = 17.74$, $p < 0.001$, $\eta^2 = 0.05$; stimulus type: $F(1, 18) = 41.47$, $p < 0.001$, $\eta^2 = 0.07$; train length: $F(4, 72) = 3.70$, $p = 0.01$, $\eta^2 = 0.02$] and significant two-way interactions [condition \times stimulus type: $F(2, 36) = 4.37$, $p = 0.02$, $\eta^2 = 0.02$; condition \times train length: $F(8, 144) = 4.53$, $p < 0.001$, $\eta^2 = 0.03$; stimulus type \times train length: $F(4, 72) = 26.05$, $p < 0.001$, $\eta^2 = 0.11$]. All those effects were qualified by a three-way interaction condition \times stimulus type \times train length: $F(8, 144) = 2.54$, $p = 0.01$, $\eta^2 = 0.03$.

Subsequently, we first tested whether in each of the three conditions the deviant ERP differed from the standard ERP (main effect of stimulus type) and whether this difference developed with increasing train length (interaction stimulus type \times train length). Both the stimulus type effect [exact: $F(1, 18) = 25.40$,



$p < 0.001$, $\eta^2 = 0.16$; wn: $F(1, 18) = 5.54$, $p = 0.03$, $\eta^2 = 0.03$; wp: $F(1, 18) = 5.76$, $p = 0.03$, $\eta^2 = 0.03$ and the interaction stimulus type \times train length [exact: $F(4, 72) = 19.25$, $p < 0.001$, $\eta^2 = 0.20$; wn: $F(4, 72) = 5.17$, $p = 0.001$, $\eta^2 = 0.10$; wp: $F(4, 72) = 4.05$, $p = 0.01$, $\eta^2 = 0.06$] were significant in each condition. Particularly, the stimulus type \times train length interactions resulted from increasingly negative MMN amplitudes of the deviant minus standard difference with train length (see **Figure 5A**).

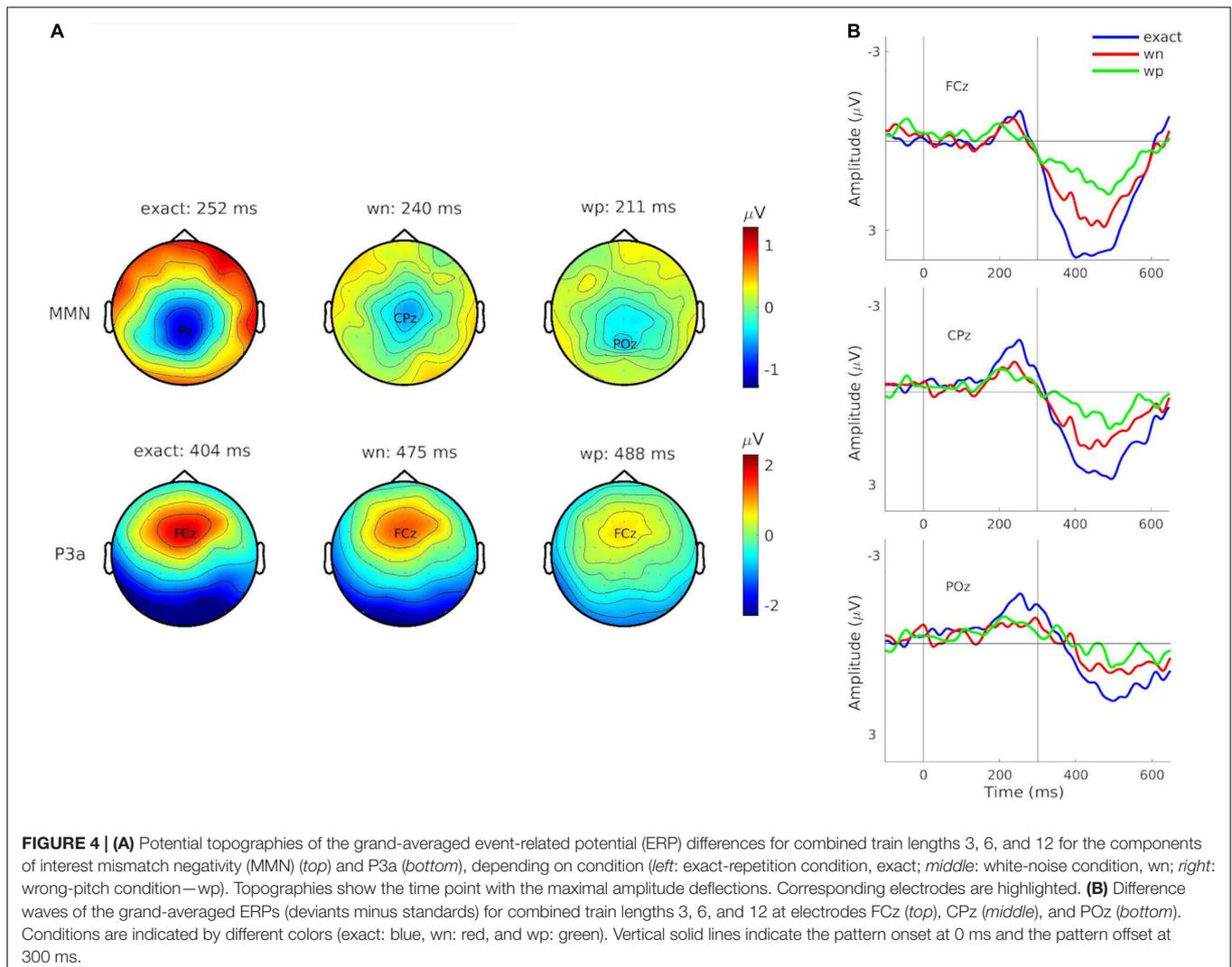
To further explore the origin of the condition \times stimulus type \times train length interaction, we analyzed deviant and standard amplitudes at the centro-parietal electrodes in separate two-way ANOVAs for possible train length effects and interactions with condition. Here, we found for the deviants a significant condition [$F(2, 36) = 12.03$, $p < 0.001$, $\eta^2 = 0.11$] and train length effect [$F(4, 72) = 20.92$, $p < 0.001$, $\eta^2 = 0.16$] and a significant condition \times train length interaction [$F(8, 144) = 4.97$, $p < 0.001$, $\eta^2 = 0.08$]. However, for standards, the two-way ANOVA revealed only a train length effect [$F(4, 72) = 8.05$, $p < 0.001$, $\eta^2 = 0.07$]. The condition effect [$F(2, 36) = 2.41$, $p = 0.10$, $\eta^2 = 0.01$] and the condition \times train length interaction [$F(8, 144) = 1.19$, $p = 0.31$, $\eta^2 = 0.02$] were not significant. This indicates a similar development of the standard amplitudes with train length independently of condition and a different development of deviant amplitudes with train length in dependence of condition. Furthermore, we observed a significant logarithmic trend for the increase in deviant negativity with increasing train length in all three conditions [exact: $F(1, 18) = 67.48$, $p < 0.001$, $\eta^2 = 0.79$; wn: $F(1, 18) = 13.22$, $p = 0.002$, $\eta^2 = 0.42$; wp: $F(1, 18) = 5.20$, $p = 0.04$,

$\eta^2 = 0.22$] (see **Table 1**). Logarithmic regression analyses revealed a steeper increase in negativity for deviants in the exact-repetition condition ($r = -0.82$; CI: -1.05 to -0.59) than in the white noise ($r = -0.28$; CI: -0.49 to -0.06) and wrong-pitch condition ($r = -0.20$; CI: -0.39 to -0.02) (see **Table 1** and **Figure 5B**).

P3a Mean Amplitudes

A repeated measures ANOVA with the factors condition (exact, wn, and wp), stimulus type (dev, stand) and train length (1, 2, 3, 6, and 12) for mean amplitudes in the P3a time window (400 to 500 ms after the stimulus onset) revealed a significant main effect of condition [$F(2, 36) = 6.47$, $p = 0.004$, $\eta^2 = 0.01$]. Furthermore, the main effect of stimulus type [$F(1, 18) = 122.86$, $p < 0.001$, $\eta^2 = 0.36$] and the main effect of train length [$F(4, 72) = 4.64$, $p = 0.002$, $\eta^2 = 0.01$] were found. The two-way interactions condition \times stimulus type [$F(1.50, 26.97) = 28.57$, $p < 0.001$, $\eta^2 = 0.07$] and train length \times stimulus type [$F(2.76, 49.70) = 38.45$, $p < 0.001$, $\eta^2 = 0.09$], as well as the interaction condition \times train length [$F(8, 144) = 2.18$, $p = 0.03$, $\eta^2 = 0.01$], were significant. However, the main effects and the two-way interactions were qualified by an additional three-way interaction of the factors condition \times stimulus type \times train length: [$F(8, 144) = 5.13$, $p < 0.001$, $\eta^2 = 0.02$].

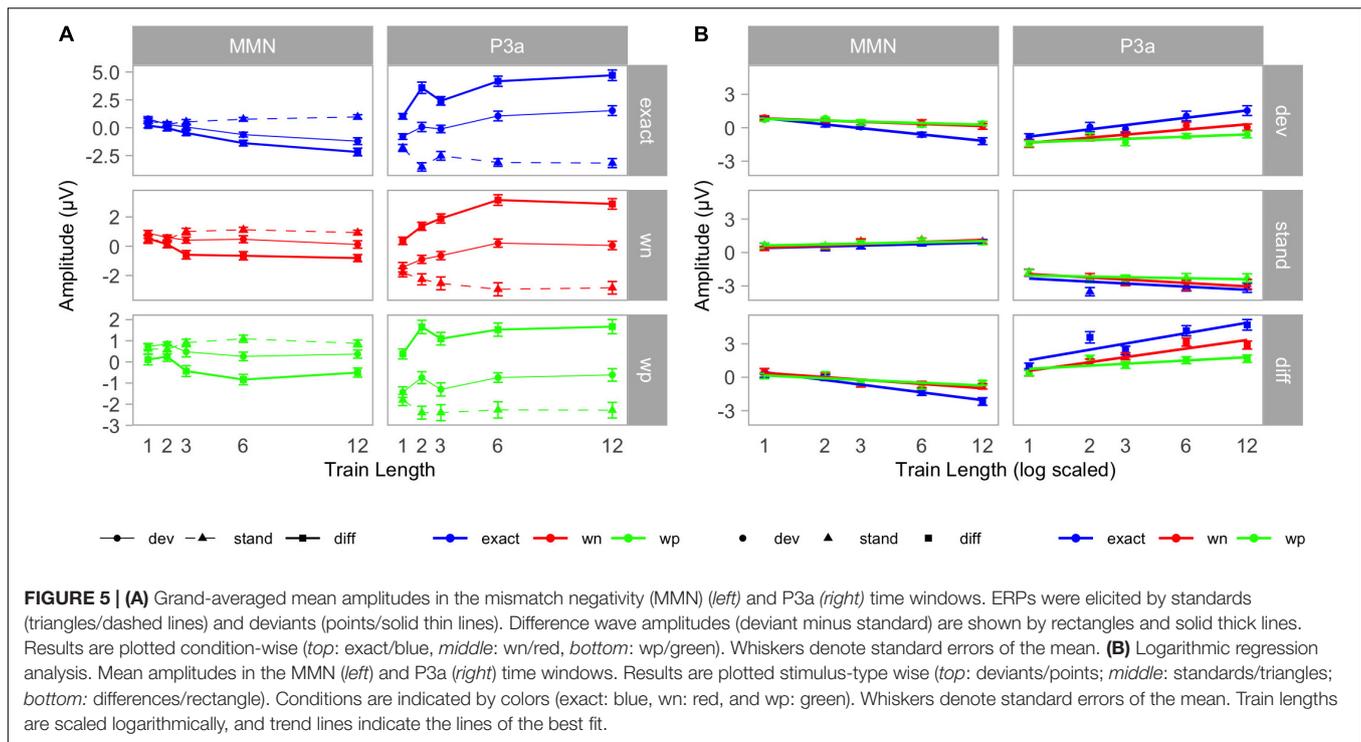
Two-way repeated measures ANOVAs with the factors stimulus type (dev, stand) and train length (1, 2, 3, 6, and 12) revealed the main effect of stimulus type [exact: $F(1, 18) = 95.95$, $p < 0.001$, $\eta^2 = 0.52$; wn: $F(1, 18) = 86.24$, $p < 0.001$, $\eta^2 = 0.31$; wp:



$F(1, 18) = 57.71, p < 0.001, \eta^2 = 0.20$) and an interaction between stimulus type and train length [exact: $F(4, 72) = 26.51, p < 0.001, \eta^2 = 0.16$; wn: $F(4, 72) = 20.31, p < 0.001, \eta^2 = 0.11$; wp: $F(4, 72) = 4.09, p = 0.005, \eta^2 = 0.04$] for each experimental condition. Particularly, the stimulus type \times train length interactions resulted from increasing positive P3a amplitudes of the deviant minus standard difference with train length (see **Figure 5A**).

To further explore the origin of the condition \times stimulus type \times train length interaction, we analyzed deviants and standards in separate two-way ANOVAs for possible train length effects and interactions with conditions. Here, we found for the deviants a significant condition [$F(2, 36) = 24.56, p < 0.001, \eta^2 = 0.15$] and train length effect [$F(4, 72) = 31.84, p < 0.001, \eta^2 = 0.16$], which was qualified by a significant condition \times train length interaction [$F(8, 144) = 3.90, p < 0.001, \eta^2 = 0.04$]. For standards, the two-way ANOVA revealed a train length effect [$F(4, 72) = 11.76, p < 0.001, \eta^2 = 0.06$] and a condition effect [$F(2, 36) = 10.07, p < 0.001, \eta^2 = 0.03$]. Also, the condition \times train length interaction [$F(8, 144) = 3.78, p < 0.001, \eta^2 = 0.02$] was significant.

Despite significant effects and interactions for deviants and standards, the effects differed between conditions in their strength as *post hoc* logarithmic trend analysis shows (see **Table 1**). We observed a significant logarithmic trend for the increase in deviant positivity with increasing train length in all three conditions [exact: $F(1, 18) = 47.90, p < 0.001, \eta^2 = 0.73$; wn: $F(1, 18) = 75.38, p < 0.001, \eta^2 = 0.81$; wp: $F(1, 18) = 6.51, p = 0.02, \eta^2 = 0.27$]. Logarithmic regression analyses revealed a steeper increase in positivity for deviants in the exact-repetition condition ($r = 0.95$; CI: 0.57–1.32) compared with the wrong-pitch condition ($r = 0.29$; CI: 0.01–0.56). Confidence intervals in the white-noise condition ($r = 0.67$; CI: 0.38–0.96) overlapped with the increase in the exact-repetition condition and the wrong-pitch condition. For standards, we did not find a significant logarithmic trend in the wrong-pitch condition [$F(1, 18) = 1.65, p = 0.22, \eta^2 = 0.08$] in contrast to the exact-repetition condition [$F(1, 18) = 11.70, p < 0.001, \eta^2 = 0.40$] and the white-noise condition [$F(1, 18) = 18.03, p < 0.001, \eta^2 = 0.50$]. The difference waves of the wrong-pitch condition also developed less steeply than in the exact-repetition condition and in the white-noise



condition as confidence intervals did not overlap (exact: $r = 1.35$; CI: 0.91–1.80, wn: $r = 1.11$; CI: 0.79–1.44, wp: $r = 0.42$; CI: 0.11–0.73) (see **Table 1** and **Figure 5B**).

DISCUSSION

In this study, we investigated the implicit memory formation for repeated auditory objects in situations in which single occurrences of the same object were subject to variability. More specifically, in three conditions, we studied how repetitions of unfamiliar short sound patterns lead to the formation of pattern-specific auditory sensory memory representations when single instances of pattern repetitions are identical (exact) or when they contain small distortions by replacing a segment of the pattern information either by white noise (wn) or by a wrong pitched segment (wp). The participants were not explicitly focusing on the pattern-repetition rule and, instead, performed the task of detecting occasionally occurring loudness changes in the auditory sequence. They performed this task with high accuracy and in the second session of the experiment with increased sensitivity and faster reaction times, likely due to familiarity and learning effects. Sensitivity of the participants to detect loudness changes was not affected by condition. However, the reaction times were significantly slower in the white-noise condition. It is possible the white-noise insertions were stronger distractors from the task, because these segments differed in sound quality from all the other segments of the sound pattern.

As expected, the occurrence of a new sound pattern elicited an MMN and a subsequent P3a; both of which appeared with their typical time course according to previous studies, using

an auditory oddball (Alho et al., 1997; Schröger and Wolff, 1998b; Roeber et al., 2003; Debener et al., 2005) and roving-standard paradigms (Cowan et al., 1993; Baldeweg et al., 2004; Bendixen and Schröger, 2008; Garrido et al., 2008). However, the central-parietal distribution of an MMN seems rather untypical, since it can usually be observed at fronto-central recording sites (Näätänen et al., 2001, 2007; Winkler, 2007). We already observed rather posterior MMN topography in a previous study, using similar stimulus material in a comparable experimental procedure (Bader et al., 2017). An explanation for this topography could be the fact that the participants attended the sound sequence (although the pattern changes themselves were task irrelevant). That is, attention might have modulated the otherwise automatic mismatch detection process, potentially allowing for the contribution of an N2b-like component. Yet, even in this case, a more central but not a posterior distribution is expected (e.g., Woods, 1992). Furthermore, the task irrelevance of the pattern changes, and the early latency of the negative peak argues rather against an N2b. Coy et al. (2021) indicate that N2b for task-relevant pattern changes occurs later and can be dissociated from MMN, since, in their data, N2b latency is modulated by the difficulty of the deviant discrimination, whereas MMN latency is not. An alternative explanation could be that the auditory task increased the distracting nature of the deviants (Schröger and Wolff, 1998a). This could evoke a prominent and early P3a in the current paradigm (compared with the more typical passive listening situation). An early emergence of a frontally distributed P3a might partly overlap with the MMN time window and shift the topography toward more posterior sites. Nevertheless, more posterior distributions of the MMN have also been reported in other studies in which the participants

TABLE 1 | Logarithmic trend and regression analysis for the main effect of train length on MMN and P3a amplitudes.

	MMN				P3a			
	<i>F</i> (1, 18)	<i>P</i>	η^2	β (upper bound - lower bound)	<i>F</i> (1, 18)	<i>P</i>	η^2	β (upper bound - lower bound)
standards								
exact	4.82	0.04	0.21	0.19 (−0.00 to 0.38)	11.70	0.00	0.40	−0.41 (−0.79 to −0.03)
wn	14.25	0.00	0.44	0.28 (0.09 to 0.46)	18.03	0.00	0.50	−0.45 (−0.84 to −0.05)
wp	4.12	0.06	0.19	0.16 (−0.03 to 0.35)	1.65	0.22	0.08	−0.14 (−0.48 to 0.21)
deviants								
exact	67.48	0.00	0.79	−0.82 (−1.05 to −0.59)	47.90	0.00	0.73	0.95 (0.57 to 1.32)
wn	13.22	0.00	0.42	−0.28 (−0.49 to −0.06)	75.38	0.00	0.81	0.67 (0.38 to 0.96)
wp	5.20	0.04	0.22	−0.20 (−0.39 to −0.02)	6.51	0.02	0.27	0.29 (0.01 to 0.56)
difference								
exact	52.36	0.00	0.74	−1.01 (−1.27 to −0.75)	70.57	0.00	0.80	1.35 (0.91 to 1.80)
wn	20.18	0.00	0.53	−0.55 (−0.82 to −0.29)	58.64	0.00	0.77	1.11 (0.79 to 1.44)
wp	7.07	0.02	0.28	−0.36 (−0.61 to −0.12)	7.04	0.02	0.28	0.42 (0.11 to 0.73)

watched a silent movie, suggesting that the complexity of the auditory stimuli (e.g., speech and action words) could also result in such atypical MMN topography, indicating activation of a more global network (Shtyrov et al., 2004; Hasting et al., 2007).

Sensory Memory Trace Formation as Indexed by MMN

The MMM implies that the regularity of repeated complex sound patterns was encoded into a predictive model, and that a change in the overall pattern was detected (Denham and Winkler, 2006; Bendixen et al., 2012; Winkler and Schröger, 2015). This was the case for exact pattern repetitions as well as patterns containing distortions of quality of one segment (white noise) or pitch. We conclude that the auditory system quickly forms pattern representations, even when distortions introduce uncertainty into the implicit learning process. Nevertheless, the amplitude of the MMM was smaller with standard variability, suggesting diminished precision of the predictive model.

Yet MMN amplitude was larger in the condition without distortions compared with the conditions with distortions in the standard patterns. This suggests that the certainty about the repetition regularity must be higher in the exact-repetition condition, which results in more pronounced deviant-related negativity. In particular, MMN amplitudes increased as a function of preceding number of standards, following a logarithmic trend in all conditions, suggesting a fast buildup of pattern representations, particularly with the initial pattern repetitions (see also Baldeweg et al., 2004; Costa-Faidella et al., 2011; Bader et al., 2017). In general, two effects caused the growth in MMN: an increasing positivity for standard ERPs and an increasing negativity for deviant ERPs in the respective MMN time window as a function of train length. The steepness, with which MMN amplitudes grew with increasing train length, was modulated by condition. This effect was mainly driven by the deviant responses, which yielded a steeper growth function in the exact repetition condition than in the conditions containing distortions. The train length effects on standard responses were

not modulated by condition, yet, already, their size was smaller than those for deviants.

While change detection was affected by whether distortions were introduced or not, it was not modulated by the type of distortion (noise or wrong pitch information). It seems to be mainly the number of intact segments within the sound patterns, guiding the fundamental process of extracting pattern identity representations as generative predictors under uncertainty, at least on this early processing stage.

The Role of Exact and Distorted Pitch Code on Evaluation Processes as Indexed by P3a and Behavior in an Active Pattern Change Detection Task

Subsequent to the MMN, a P3a component with typical fronto-central distribution was elicited, and systematic repetition-related modulations of amplitudes were found in all the conditions (see also Bendixen et al., 2007; Horvath et al., 2008; Bader et al., 2017). P3a amplitude increased logarithmically as a function of the number of preceding standard stimuli in all the conditions. This is congruent with studies showing that the P300 amplitude for task-irrelevant deviants is increased, if they occur with lower probability (Squires et al., 1975; Katayama and Polich, 1996) since, in our study, decreased local deviant probability (resulting from longer train lengths) led to an increase of P3a. Overall, P3a magnitude is associated with the degree of novelty and constitutes a marker of the evaluative processing of the contextual novelty (Friedman et al., 2001; Bendixen and Schröger, 2008). Even though attention was focused on a rule-independent task, deviant stimuli likely captured involuntarily attention and were evaluated on the basis of their underlying pattern structure in all conditions.

The growth of P3a as a function of train length resulted from both an increasing positivity for deviant ERPs and an increasing negativity for standard ERPs in the respective P3a time window. For both stimulus types, the growth function was modulated by condition. At the level of the difference waveform, the steepest logarithmic increase in P3a amplitudes with train length was found in the exact-repetition condition. The logarithmic

growth of P3a amplitudes in the wrong-pitch condition was distinctly less steep compared with the white noise and the exact-repetition condition. Thus, the growth of P3a activity with increasing number of preceding standards seems to be lowest in the condition with wrong pitch information. Consequently, pattern changes in the context of exact repetitions and white noise distortions might need fewer standard presentations to evaluate deviants as potential novelties or targets and to trigger a call for attention of equal strength than pattern changes occurring among repetitions with wrong pitch distortions. Condition differences between P3a growths were partly due to the processing of the standard events. In the wrong-pitch condition, standard responses in the P3a time window did not show a logarithmic repetition-related modulation as opposed to the exact repetition and the white-noise-distortion condition. The amplitude decreases with the number of repetitions seems relatively flat, but confidence intervals overlapped with the exact repetition and the white-noise-distortion condition. Overall, the processing of deviating events had a stronger contribution to the condition differences on the P3a level (similar to what we found for the MMN). In the exact-repetition condition, the logarithmic trend of systematic amplitude modulations was much steeper than in the wrong-pitch condition.

Despite the pronounced differences in P3a amplitude, estimated component latencies were not affected by our experimental manipulation. The time needed to internally evaluate pattern identity seems not higher with noise or wrong pitch insertions. Here, the auditory system seems equally quick in classifying the stimuli, in evaluating the novelty of the deviance, and in automatically orienting attention toward the task-irrelevant novel patterns (Polich, 2007; Horvath et al., 2008; Wetzel et al., 2013; Winkler and Schröger, 2015).

The behavioral performance in the active pattern change detection task showed highest accuracy in the exact-repetition condition, less accuracy in the white-noise condition, and least accuracy in the wrong pitch segment condition. The development of the hit rates as a function of train length also followed a logarithmic trend in all conditions with the hit rates in the exact-repetition condition, developing quickly toward a ceiling effect and most slowly in the wrong-pitch condition, mirroring the pattern of results at the level of P3a.

Effects of the Different Types of Distortion at the Levels of Sensory Memory Formation and Contextual Stimulus Evaluation

Effects in the MMN time range did not distinguish between the two types of distortions that we introduced in the sequences. In both cases, when a pattern segment was occasionally replaced by white noise or by a differently pitched segment, the typical negativity observed for deviant responses grew less steeply compared with the condition with exact pattern repetitions. As the portion of matching pattern segments was similar for both conditions (five out of six), one could speculate that, at this early processing stage, the strength of regularity encoding mainly depends on the probability of matching information in

spectrotemporal space. A previous study revealed that object representations can be retrieved from repetitions embedded in different backgrounds and suggested the correlation of input spectrograms between several non-perfect repetitions as a potential mechanism to achieve this (McDermott et al., 2011).

In contrast to the MMN findings, the type of distortion (white noise or wrong pitch information) seems to have a greater impact on evaluation processes at the stage of P3a. Here, we observed a clear disadvantage for the wrong pitch compared with the exact-repetition condition, whereas the response pattern in the white-noise condition resembled that of the exact-repetition condition. This suggests that the higher cognitive evaluation process and the attentional switch toward a deviant pattern are not substantially impaired in the white-noise condition. This could be explained by processes related to stream segregation or perceptual filling-in. Firstly, the white noise segments differed substantially from the other five segments in their spectral composition, leading to a vastly different timbre percept. This could have assisted the segregation of pattern and distracter information, allowing the later stages of evaluating a newly incoming pattern to be less disturbed. Secondly, white noise insertions could even lead to a partial restoration of the perceptual continuity of an intact standard pattern (Warren et al., 1997, 1988), particularly since the pitch forming spectral components of the replaced pattern segment can physically be found in the 50-ms white noise segment. This could explain similar processing of standards and deviants in the exact and the white-noise condition. The processing of wrong pitch information might have prevented segregation and restoration processes. Here, the similar timbre of the distorted segment promoted the percept of a continuous sound pattern, and a segregation mechanism cannot take an effect. Also, the wrong pitch information is simply misleading, and there is no option for a perceptual restoration of erased or ambiguous information. This might have led to the greater disadvantage for the wrong-pitch condition and is compatible with explanations of failed segregation in informational masking studies (Kidd et al., 2008).

Overall, our two types of manipulations might remind of energetic and informational masking. During energetic masking, the processing of a sensory event is degraded typically by a noise masker, where, mainly, the energy relation between the signal and the masker determines their separability and the amount of difficulty in perceiving the signal (Muller-Gass et al., 2001; Darwin, 2008; Moore, 2008; Wilson et al., 2012). During informational masking, it is, additionally, the similarity (either the acoustic or the semantic similarity) between the target and masker sound that deteriorates segregation and perception of the signal sound, such as when a speech signal is masked by speech (Dirks and Bower, 1969; Durlach et al., 2003; Scott et al., 2004; Gutschalk et al., 2008; Kidd et al., 2008; Mattys et al., 2012). In masking studies, it has been found that irrelevant information is easier to suppress when energetic masking is dominant, thus facilitating the processing of relevant sensory information (Lidestam et al., 2014). In contrast, when informational masking is dominant, a widespread general attentional network is activated to distinguish distracter and target information (Szalárdy et al., 2019). In our study, the wrong

pitch segment is due to its spectral profile acoustically highly similar to the pattern elements themselves. The uncertainty about which are the five relevant (intact) segments of the tonal pattern is maximally high, and segregation between the true pattern and the replaced distracting element will not be as easy as in the white-noise condition. In that sense, replacing valid pattern information by white noise or by wrong pitch information could lead to distinct forms of interference that might be resolved at different processing levels and even distinguishable neural circuits (Scott et al., 2004).

CONCLUSION

In the present study, we demonstrated that the auditory system is able to form pattern representations and predictions even in the context of uncertainty. MMN and P3a were elicited in response to deviants in all conditions at similar latency estimations. Independent of distorted segments, an implicit and automatic buildup of regularity representations and deviance detection, with a following call for attention toward the stimuli, can be assumed. However, we found a general advantage for the exact-repetition condition over the two distorted repetition conditions. This is shown by steeper logarithmic amplitude changes on MMN and the P3a level, with increasing number of repetitions, as well as in the behavioral performance in the active pattern change detection task. The processing on the MMN level does not seem to differentiate between the qualities of the distortions but reacts to general (mis-)matching statistics between the sound patterns. At the level of P3a, we observed an influence of the type of distortion. On this processing stage, white-noise distortions may not have impeded stimulus processing to such an extent as wrong pitch distortions did. This is evidenced by the distinct degradation of the P3a, which also goes along with the behavioral findings in the active pattern change detection task. Additionally, our results revealed that deviant processing is more affected by our experimental manipulations compared with standard processing.

To sum up, our findings indicate that the auditory system is able to quickly extract regularities and generate reliable predictions, even when the to-be-extracted patterns contain distortions. However, higher cognitive processing and the involvement of an attentional network might give the basis for the subtler evaluation of the acoustic input with regard to the type of distortion. The segregation between informative pattern segments and the distorting element, as well as the accessibility of a possible interpolation mechanism (like the one discussed for

the continuity illusion), might explain the facilitated processing of white noise compared with wrong-pitch distortions on later processing stages that govern the elicitation of P3a and guide behavioral responses.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

This studies involving human participants were reviewed and approved by the Ethics Committee of the University of Leipzig, Germany. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SG and ES: conceptualization, supervision, and resources. MB: data curation, visualization, and writing—original draft. MB and SG: formal analysis, investigation, software, and project administration. SG: funding acquisition. SG, ES, and MB: methodology, writing, review, and editing, and validation. All authors contributed to the article and approved the submitted version.

FUNDING

The study was funded by the German Research Foundation (www.dfg.de) with the project number GR 3412/2-1. Furthermore, we acknowledge support from the German Research Foundation (DFG) and the Leipzig University within the program of Open Access Publishing. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We wish to thank Nina Coy for helping with the data acquisition and Robert P. O'Shea for commenting helpfully on the previous draft of the manuscript.

REFERENCES

- Alho, K., Escera, C., Diaz, R., Yago, E., and Serra, J. M. (1997). Effects of involuntary auditory attention on visual task performance and brain activity. *Neuroreport* 8, 3233–3237. doi: 10.1097/00001756-199710200-00010
- Bader, M., Schröger, E., and Grimm, S. (2017). How regularity representations of short sound patterns that are based on relative or absolute pitch information establish over time: an EEG study. *PLoS One* 12:e0176981. doi: 10.1371/journal.pone.0176981
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384. doi: 10.3758/bf03192707
- Baldeweg, T., Klugman, A., Gruzelier, J., and Hirsch, S. R. (2004). Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophr. Res.* 69, 203–217. doi: 10.1016/j.schres.2003.09.009
- Bendixen, A., Roeber, U., and Schröger, E. (2007). Regularity extraction and application in dynamic auditory stimulus sequences. *J. Cogn. Neurosci.* 19, 1664–1677. doi: 10.1162/jocn.2007.19.10.1664

- Bendixen, A., SanMiguel, I., and Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: a review. *Int. J. Psychophysiol.* 83, 120–131. doi: 10.1016/j.jpsycho.2011.08.003
- Bendixen, A., Scharinger, M., Strauss, A., and Obleser, J. (2014). Prediction in the service of comprehension: modulated early brain responses to omitted speech segments. *Cortex* 53, 9–26. doi: 10.1016/j.cortex.2014.01.001
- Bendixen, A., and Schröger, E. (2008). Memory trace formation for abstract auditory features and its consequences in different attentional contexts. *Biol. Psychol.* 78, 231–241. doi: 10.1016/j.biopsycho.2008.03.005
- Berti, S., Roeber, U., and Schröger, E. (2004). Bottom-up influences on working memory: behavioral and electrophysiological distraction varies with distractor strength. *Exp. Psychol.* 51, 249–257. doi: 10.1027/1618-3169.51.4.249
- Brattico, E., Tervaniemi, M., Näätänen, R., and Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Res.* 1117, 162–174. doi: 10.1016/j.brainres.2006.08.023
- Brown, G. S., and White, K. G. (2005). The optimal correction for estimating extreme discriminability. *Behav. Res. Methods* 37, 436–449.
- Cooke, M., Garcia Lecumberri, M. L., and Barker, J. (2008). The foreign language cocktail party problem: energetic and informational masking effects in non-native speech perception. *J. Acoust. Soc. Am.* 123, 414–427. doi: 10.1121/1.2804952
- Costa-Faidella, J., Baldeweg, T., Grimm, S., and Escera, C. (2011). Interactions between "What" and "When" in the auditory system: temporal predictability enhances repetition suppression. *J. Neurosci.* 31, 18590–18597. doi: 10.1523/Jneurosci.2599-11.2011
- Cowan, N., Winkler, I., Feder, W., and Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (Erp). *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 909–921. doi: 10.1037/0278-7393.19.4.909
- Coy, N., Bader, M., Schröger, E., and Grimm, S. (2021). Change detection of auditory tonal patterns defined by absolute versus relative pitch information. A combined behavioural and EEG study. *PLoS One* 16:e0247495. doi: 10.1371/journal.pone.0247495
- Darwin, C. J. (2008). Listening to speech in the presence of other sounds. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1011–1021. doi: 10.1098/rstb.2007.2156
- De Weerd, P. (2006). Perceptual filling-in: more than the eye can see. *Prog. Brain Res.* 154, 227–245. doi: 10.1016/S0079-6123(06)54012-9
- Debener, S., Makeig, S., Delorme, A., and Engel, A. K. (2005). What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Cogn. Brain Res.* 22, 309–321. doi: 10.1016/j.cogbrainres.2004.09.006
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Denham, S. L., and Winkler, I. (2006). The role of predictive models in the formation of auditory streams. *J. Physiol. Paris* 100, 154–170. doi: 10.1016/j.jphysparis.2006.09.012
- DeWitt, L. A., and Samuel, A. G. (1990). The role of knowledge-based expectations in music perception: evidence from musical restoration. *J. Exp. Psychol. Gen.* 119, 123–144. doi: 10.1037//0096-3445.119.2.123
- Dirks, D. D., and Bower, D. R. (1969). Masking effects of speech competing messages. *J. Speech Hear. Res.* 12, 229–245. doi: 10.1044/jshr.1202.229
- Durlach, N. I., Mason, C. R., Kidd, G. Jr., Arbogast, T. L., Colburn, H. S., and Shinn-Cunningham, B. G. (2003). Note on informational masking (L). *J. Acoust. Soc. Am.* 113, 2984–2987. doi: 10.1121/1.1570435
- Escera, C., Alho, K., Schröger, E., and Winkler, I. W. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *J. Audiol. Otol.* 5, 151–166.
- Friedman, D., Cycowicz, Y. M., and Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci. Biobehav. Rev.* 25, 355–373. doi: 10.1016/S0149-7634(01)00019-7
- Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., Baldeweg, T., and 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., and Friston, K. J. (2009). Repetition suppression and plasticity in the human brain. *Neuroimage* 48, 269–279. doi: 10.1016/j.neuroimage.2009.06.034
- Gutschalk, A., Micheyl, C., and Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biol.* 6:e138. doi: 10.1371/journal.pbio.0060138
- Hasting, A. S., Kotz, S. A., and Friederici, A. D. (2007). Setting the stage for automatic syntax processing: the mismatch negativity as an indicator of syntactic priming. *J. Cogn. Neurosci.* 19, 386–400. doi: 10.1162/jocn.2007.19.3.386
- Hautus, M., and Lee, A. (2006). Estimating sensitivity and bias in a yes/no task. *Br. J. Math. Stat. Psychol.* 59, 257–273. doi: 10.1348/000711005X65753
- Horvath, J., Winkler, I., and Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biol. Psychol.* 79, 139–147. doi: 10.1016/j.biopsycho.2008.04.001
- Katayama, J., and Polich, J. (1996). P300, probability, and the three-tone paradigm. *Evoked Potent. Electroencephalogr. Clin. Neurophysiol.* 100, 555–562. doi: 10.1016/S0168-5597(96)95171-0
- Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. I. (2008). "Informational Masking" in *Auditory Perception of Sound Sources*, eds W. A. Yost, A. N. Popper, and R. R. Fay (Boston, MA: Springer), 143–189.
- Kiesel, A., Miller, J., Jolicoeur, P., and Brisson, B. (2008). Measurement of ERP latency differences: a comparison of single-participant and jackknife-based scoring methods. *Psychophysiology* 45, 250–274. doi: 10.1111/j.1469-8986.2007.00618.x
- Kozou, H., Kujala, T., Shtyrov, Y., Toppila, E., Starck, J., Alku, P., et al. (2005). The effect of different noise types on the speech and non-speech elicited mismatch negativity. *Hear. Res.* 199, 31–39. doi: 10.1016/j.heares.2004.07.010
- Lidestam, B., Holgersson, J., and Moradi, S. (2014). Comparison of informational vs. energetic masking effects on speechreading performance. *Front. Psychol.* 5:639. doi: 10.3389/fpsyg.2014.00639
- Macmillan, N. A., and Creelman, C. D. (2004). *Detection Theory: A User's Guide*. London: Psychology press.
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024
- Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006
- McDermott, J. H., Wroblewski, D., and Oxenham, A. J. (2011). Recovering sound sources from embedded repetition. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1188–1193. doi: 10.1073/pnas.1004765108
- Micheyl, C., Carlyon, R. P., Shtyrov, Y., Hauk, O., Dodson, T., and Pullvermüller, F. (2003). The neurophysiological basis of the auditory continuity illusion: a mismatch negativity study. *J. Cogn. Neurosci.* 15, 747–758. doi: 10.1162/089892903322307456
- Miller, J., Patterson, T., and Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology* 35, 99–115.
- Moore, B. C. (2008). Basic auditory processes involved in the analysis of speech sounds. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* 363, 947–963. doi: 10.1098/rstb.2007.2152
- Muhler, R., and von Specht, H. (1999). Sorted averaging—principle and application to auditory brainstem responses. *Scand. Audiol.* 28, 145–149. doi: 10.1080/010503999424716
- Muller-Gass, A., Marcoux, A., Logan, J., and Campbell, K. B. (2001). The intensity of masking noise affects the mismatch negativity to speech sounds in human subjects. *Neurosci. Lett.* 299, 197–200. doi: 10.1016/S0304-3940(01)01508-7
- Näätänen, R., Paavilainen, P., Rinne, T., and 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
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., and Winkler, I. (2001). 'Primitive intelligence' in the auditory cortex. *Trends Neurosci.* 24, 283–288. doi: 10.1016/S0166-2236(00)01790-2
- Niemitalo-Haapala, E., Haapala, S., Jansson-Verkasalo, E., and Kujala, T. (2015). Background noise degrades central auditory processing in toddlers. *Ear. Hear.* 36, E342–E351. doi: 10.1097/Aud.0000000000000192
- Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:156869. doi: 10.1155/2011/156869

- Paavilainen, P., Karlsson, M. L., Reinikainen, K., and Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalogr. Clin. Neurophysiol.* 73, 129–141. doi: 10.1016/0013-4694(89)90192-2
- Perruchet, P., and Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *TiCS* 10, 233–238. doi: 10.1016/j.tics.2006.03.006
- 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
- Rahne, T., von Specht, H., and Muhler, R. (2008). Sorted averaging—application to auditory event-related responses. *J. Neurosci. Methods* 172, 74–78. doi: 10.1016/j.jneumeth.2008.04.006
- Ramachandran, V. S. (1992). Filling in the blind spot. *Nature* 356, 115–115. doi: 10.1038/356115a0
- Riecke, L., van Opstal, A. J., Goebel, R., and Formisano, E. (2007). Hearing illusory sounds in noise: sensory-perceptual transformations in primary auditory cortex. *J. Neurosci.* 27, 12684–12689. doi: 10.1523/Jneurosci.2713-07.2007
- Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., et al. (1992). Event-related potentials to repetition and change of auditory-stimuli. *Electroencephalogr. Clin. Neurophysiol.* 83, 306–321. doi: 10.1016/0013-4694(92)90090-5
- Roeber, U., Berti, S., and Schröger, E. (2003). Auditory distraction with different presentation rates: an event-related potential and behavioral study. *Clin. Neurophysiol.* 114, 341–349. doi: 10.1016/S1388-2457(02)00377-2
- Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 62, 437–448. doi: 10.1016/0168-5597(85)90054-1
- Schröger, E. (1997). On the detection of auditory deviations: a pre-attentive activation model. *Psychophysiology* 34, 245–257. doi: 10.1111/j.1469-8986.1997.tb02395.x
- Schröger, E., and Wolff, C. (1998a). Attentional orienting and reorienting is indicated by human event-related brain potentials. *Neuroreport* 9, 3355–3358. doi: 10.1097/00001756-199810260-00003
- Schröger, E., and Wolff, C. (1998b). Behavioral and electrophysiological effects of task-irrelevant sound change: a new distraction paradigm. *Cogn. Brain Res.* 7, 71–78.
- Scott, S. K., Rosen, S., Wickham, L., and Wise, R. J. S. (2004). A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. *J. Acoust. Soc. Am.* 115, 813–821. doi: 10.1121/1.1639336
- Shahin, A. J., Bishop, C. W., and Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage* 44, 1133–1143. doi: 10.1016/j.neuroimage.2008.09.045
- Shinn-Cunningham, B. G., and Wang, D. J. (2008). Influences of auditory object formation on phonemic restoration. *J. Acoust. Soc. Am.* 123, 295–301. doi: 10.1121/1.2804701
- Shtyrov, Y., Hauk, O., and Pulvermüller, F. (2004). Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. *Eur. J. Neurosci.* 19, 1083–1092. doi: 10.1111/j.1460-9568.2004.03126.x
- Spillmann, L., Otte, T., Hamburger, K., and Magnussen, S. (2006). Perceptual filling-in from the edge of the blind spot. *Vis. Res.* 46, 4252–4257. doi: 10.1016/j.visres.2006.08.033
- Squires, N. K., Squires, K. C., and Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroen. Clin. Neurol.* 38, 387–401. doi: 10.1016/0013-4694(75)90263-1
- Szalárdy, O., Tóth, B., Farkas, D., György, E., and Winkler, I. (2019). Neuronal correlates of informational and energetic masking in the human brain in a multi-talker situation. *Front. Psychol.* 10:786. doi: 10.3389/fpsyg.2019.00786
- Tervaniemi, M., Schröger, E., Saher, M., and Näätänen, R. (2000). Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans - a mismatch negativity study. *Neurosci. Lett.* 290, 66–70. doi: 10.1016/S0304-3940(00)01290-8
- Uhrig, S., Arndt, S., Möller, S., and Voigt-Antons, J. N. (2017). Perceptual references for independent dimensions of speech quality as measured by electroencephalography. *Q. User Exp.* 2:10. doi: 10.1007/s41233-017-0011-8
- Urbach, T. P., and Kutas, M. (2006). Interpreting event-related brain potential (ERP) distributions: implications of baseline potentials and variability with application to amplitude normalization by vector scaling. *Biol. Psychol.* 72, 333–343. doi: 10.1016/j.biopsycho.2005.11.012
- Wacongne, C., Changeux, J. P., and Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* 32, 3665–3678. doi: 10.1523/JNEUROSCI.5003-11.2012
- Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., and Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence (vol 20, pg 188, 2010). *Learn. Individ. Diff.* 20, 705–705. doi: 10.1016/j.lindif.2010.10.001
- Walls, G. L. (1954). The filling-in process. *Optom. Vis. Sci.* 31, 329–341.
- Warren, R. M., Hainsworth, K. R., Brubaker, B. S., Bashford, J. A., and Healy, E. W. (1997). Spectral restoration of speech: intelligibility is increased by inserting noise in spectral gaps. *Percept. Psychophys.* 59, 275–283. doi: 10.3758/Bf03211895
- Warren, R. M., and Warren, R. P. (1970). Auditory illusions and confusions. *Sci. Am.* 223, 30–36. doi: 10.1038/scientificamerican1270-30
- Warren, R. M., Wrightson, J. M., and Puretz, J. (1988). Illusory continuity of tonal and infratonal periodic sounds. *J. Acoust. Soc. Am.* 84, 1338–1342. doi: 10.1121/1.396632
- Wetzel, N., and Schröger, E. (2007). Modulation of involuntary attention by the duration of novel and pitch deviant sounds in children and adolescents. *Biol. Psychol.* 75, 24–31. doi: 10.1016/j.biopsycho.2006.10.006
- Wetzel, N., Schröger, E., and Widmann, A. (2013). The dissociation between the P 3a event-related potential and behavioral distraction. *Psychophysiology* 50, 920–930. doi: 10.1111/psyp.12072
- Widmann, A., Schröger, E., and Maess, B. (2015). Digital filter design for electrophysiological data—a practical approach. *J. Neurosci. Methods* 250, 34–46. doi: 10.1016/j.jneumeth.2014.08.002
- Wilson, R. H., Trivette, C. P., Williams, D. A., and Watts, K. L. (2012). The effects of energetic and informational masking on the Words-in-Noise Test (WIN). *J. Am. Acad. Audiol.* 23, 522–533. doi: 10.3766/jaaa.23.7.4
- Winkler, I. (2007). Interpreting the mismatch negativity. *J. Psychophysiol.* 21, 147–163. doi: 10.1027/0269-8803.21.3.147
- Winkler, I., Cowan, N., Csepe, V., Czigler, I., and Näätänen, R. (1996). Interactions between transient and long-term auditory memory as reflected by the mismatch negativity. *J. Cogn. Neurosci.* 8, 403–415. doi: 10.1162/jocn.1996.8.5.403
- Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., and Näätänen, R. (1990). The effect of small variation of the frequent auditory stimulus on the event-related brain potential to the infrequent stimulus. *Psychophysiology* 27, 228–235. doi: 10.1111/j.1469-8986.1990.tb00374.x
- Winkler, I., and Schröger, E. (2015). Auditory perceptual objects as generative models: setting the stage for communication by sound. *Brain Lang.* 148, 1–22. doi: 10.1016/j.bandl.2015.05.003
- Woods, D. L. (1992). Auditory selective attention in middle-aged and elderly subjects: an event-related brain potential study. *Electroencephalogr. Clin. Neurophysiol.* 84, 456–468. doi: 10.1016/0168-5597(92)90033-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Bader, Schröger and Grimm. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.