

That sounds awful! Does sound unpleasantness modulate the mismatch negativity and its habituation?

Hanna Ringer^{1,2}  | Sarah Alica Rösch^{1,2,3}  | Urte Roeber¹  | Julia Deller^{1,4}  |
 Carles Escera^{5,6,7}  | Sabine Grimm⁸ 

¹Wilhelm Wundt Institute for Psychology, Leipzig University, Leipzig, Germany

²International Max Planck Research School on Neuroscience of Communication, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

³Integrated Research and Treatment Center (IFB) Adiposity Diseases, Behavioral Medicine Research Unit, Leipzig University Medical Center, Leipzig, Germany

⁴Department of Psychosomatic Medicine and Psychotherapy, University of Leipzig, Leipzig, Germany

⁵Brainlab – Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, Faculty of Psychology, University of Barcelona, Barcelona, Spain

⁶Institute of Neurosciences, University of Barcelona, Barcelona, Spain

⁷Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain

⁸Physics of Cognition Lab, Institute of Physics, Chemnitz University of Technology, Chemnitz, Germany

Correspondence

Sabine Grimm, Physics of Cognition Lab, Institute of Physics, Chemnitz University of Technology, Reichenhainer Straße 70, Physikgebäude (PhG), 09126 Chemnitz, Germany.
 Email: sabine.grimm@physik.tu-chemnitz.de

Funding information

Institució Catalana de Recerca i Estudis Avançats; Stiftung Begabtenförderung Cusanuswerk

Abstract

There are sounds that most people perceive as highly unpleasant, for instance, the sound of rubbing pieces of polystyrene together. Previous research showed larger physiological and neural responses for such aversive compared to neutral sounds. Hitherto, it remains unclear whether habituation, i.e., diminished responses to repeated stimulus presentation, which is typically reported for neutral sounds, occurs to the same extent for aversive stimuli. We measured the mismatch negativity (MMN) in response to rare occurrences of aversive or neutral deviant sounds within an auditory oddball sequence in 24 healthy participants, while they performed a demanding visual distractor task. Deviants occurred as single events (i.e., between two standards) or as double deviants (i.e., repeating the identical deviant sound in two consecutive trials). All deviants elicited a clear MMN, and amplitudes were larger for aversive than for neutral deviants (irrespective of their position within a deviant pair). This supports the claim of preattentive emotion evaluation during early auditory processing. In contrast to our expectations, MMN amplitudes did not show habituation, but increased in response to deviant repetition—similarly for aversive and neutral deviants. A more fine-grained analysis of individual MMN amplitudes in relation to individual

Hanna Ringer and Sarah Alica Rösch shared first authorship.

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Psychophysiology* published by Wiley Periodicals LLC on behalf of Society for Psychophysiological Research.

arousal and valence ratings of each sound item revealed that stimulus-specific MMN amplitudes were best predicted by the interaction of deviant position and perceived arousal, but not by valence. Deviants with perceived higher arousal elicited larger MMN amplitudes only at the first deviant position, indicating that the MMN reflects preattentive processing of the emotional content of sounds.

KEY WORDS

arousal, aversive sounds, emotion, event-related potential, mismatch negativity

1 | INTRODUCTION

Human experience and behavior are inextricably linked to emotions. Emotions are widely assumed to govern the preparation of appropriate behavioral reactions in a given situation, thus rendering fast and effective emotional evaluation evolutionarily adaptive (Bradley & Lang, 2002). Across sensory modalities, emotional stimuli have been found to evoke a variety of characteristic responses, including physiological reactions, behaviors, and feelings (Bradley & Lang, 2002). In the auditory domain, unpleasant sounds (such as scraping sounds, animal cries, or angry voices) are known to elicit an enlarged startle reflex, increased skin conductance, enhanced heart rate deceleration, and stronger muscle activity (Bradley & Lang, 2000). Furthermore, unpleasant sounds are recalled more readily compared to neutral or pleasant sounds (Bradley & Lang, 2000). Among unpleasant auditory experiences, certain sounds are perceived as particularly discomforting by most people, such as, for instance, the sound of fingernails scraping across a chalkboard or rubbing two pieces of polystyrene together (Cox, 2008; Halpern et al., 1986; Schweiger Gallo et al., 2017). The discomfort induced by these, henceforth called aversive sounds, has been related to specific acoustic properties such as a high spectral frequency and rather little temporal modulation (Kumar et al., 2008). Such aversive sounds elicit a distinctive pattern of heart rate responses compared to other unpleasant, pleasant and neutral sounds (Schweiger Gallo et al., 2017). On the neural level, they activate the amygdala, reflecting the sound's association with negative emotions, and are (in comparison to neutral sounds) accompanied by higher activity in auditory cortex regions (Kumar et al., 2012). In the current study, we use electroencephalography (EEG) to compare brain responses and their habituation characteristics between such aversive sounds and sounds with neutral valence that are embedded as infrequent "deviant" sounds in an oddball paradigm while participants do not pay attention to the sound input.

The emotional valence of sounds is very likely evaluated automatically during early processing steps (e.g., Paulmann & Kotz, 2008; Schirmer et al., 2005). From other research, we know that the auditory system automatically detects and exploits regularities in the auditory environment. Specifically, the mismatch negativity (MMN) as a component of the auditory event-related potential (ERP) is often used as a measure to tap into this automatic, preattentive auditory processing (for reviews see, e.g., Escera & Malmierca, 2014; Näätänen et al., 2007; Schröger, 2005). MMN is typically studied in auditory oddball paradigms, which present participants with a sequence of frequent "standard" sounds randomly interspersed with infrequent, unexpected "deviant" sounds that differ from the standards in terms of some acoustic feature. The MMN is a characteristic frontocentral negativity that is observed in the difference of the deviant minus standard ERP between 100 ms and 250 ms relative to deviance onset, often accompanied by a polarity reversal at bilateral mastoids (Näätänen et al., 1978, 2007; Schröger, 2005). MMN has been argued to result from a mismatch between anticipated and actual auditory input, with the anticipated input stemming from an internal model of the environment that was extracted from the regularities in stimulation and used to predict upcoming sensory events (Näätänen et al., 2007; Schröger, 2005). We use MMN as a tool to study the processing of aversive and neutral sounds that deviate from the regular auditory environment.

As supported by various studies, the magnitude of the MMN elicited by a deviant is sensitive to the affective content of the eliciting sound. Most of these studies used voices (that pronounced a specific syllable) as stimuli that were arranged in typical oddball sequences, with deviants differing from the standards with respect to their emotional valence in pronunciation. For example, the MMN amplitudes were enhanced for angry compared to neutral deviants and correlated with physiological arousal in terms of heart rate acceleration, which indicated that the MMN amplitude is modulated by the deviants' (emotional) relevance for the listener (Schirmer & Escoffier, 2010). Chen et al. (2016) reported

MMN amplitudes in response to fearful deviant syllables within a sequence of happy standard syllables during both wakefulness and all stages of sleep, whereas non-vocal control stimuli elicited MMN only in wakefulness and sleep stage 3, which corroborates the proposed automaticity of an emotional change detection reflected by the MMN. Furthermore, the increased MMN amplitudes for deviants that differ from the standards in terms of their emotional valence occur in particular for certain emotions: When comparing angry and happy deviants among neutral standards, angry deviants consistently evoked a larger MMN than happy deviants even in the absence of attention to the stimulation, which indicates a processing bias towards negative emotional information (Chen et al., 2018). Likewise, disgusted (compared to happy) deviants in a sequence of neutral standards were associated with activation of the anterior insula, a brain region implicated in negative emotional experience (Chen et al., 2014). The finding that MMN amplitudes were increased and MMN latencies were decreased specifically for fearful (compared to happy, sad and neutral) voices further supports a susceptibility of the MMN to processing negative emotional content and suggests that emotional stimuli with a special relevance from an evolutionary perspective are processed preferentially (Carminati et al., 2018). Notably, the authors found that the amplitudes were also largest for fearful voices during the subsequent P3a time window (Carminati et al., 2018). P3a is often observed in response to deviant or novel sounds in auditory oddball paradigms and is commonly associated with an involuntary shift of attention towards unexpected, distracting stimuli (see, e.g., Escera et al., 1998; Polich, 2007).

To the best of our knowledge, only Czigler et al. (2007) examined electrophysiological responses to aversive (compared to neutral) environmental sounds other than voices presented as deviants. Participants were instructed to detect targets (differing in pitch) within a stream of pure tones that also contained the deviants. While both aversive and neutral deviants evoked a broadly distributed negativity between around 150 ms and 250 ms after stimulus onset, amplitudes were significantly larger for aversive than for neutral deviants (Czigler et al., 2007).

The studies cited so far provide evidence that an initial evaluation of a sound's emotional content takes place during early auditory processing and tends to be biased in favor of negative emotional information. However, certain aspects regarding the processing of aversive sounds have not yet been addressed in the literature. Specifically, it remains unclear which of the orthogonal dimensions of valence and arousal, commonly used to classify emotions (Bradley & Lang, 2000; Posner

et al., 2005; Russell, 1980), drives the preferential processing of aversive stimuli: a sound's unpleasantness (i.e., its valence) or rather the associated arousal. We address this issue by taking into account our participants' individual valence and arousal ratings of the stimuli we presented.

Besides the emotional valence, the repetition of a deviant remarkably influences MMN magnitude. EEG studies have typically shown diminished MMN in response to repeated deviants, that is, deviants directly following another deviant (e.g., Müller & Schröger, 2007; Müller et al., 2005a, 2005b; Sams et al., 1984). This effect is presumably related to the well-established phenomenon of habituation: a decrease in neural response to a repeatedly presented stimulus. Habituation enables effective processing of novel and potentially more relevant sensory events (Müller et al., 2005b). Importantly, habituation of the MMN magnitude was only observed when the two deviants that occurred in a row carried the same stimulus feature (e.g., frequency; Nousak et al., 1996), and when they were grouped into the same auditory stream (Müller et al., 2005a) or tone pair (Müller & Schröger, 2007). Critically, the MMN amplitude reduction was maximal in response to a repeated deviant that was identical to the first one (Müller et al., 2005b). A similar effect was reported for the subsequent P3a, which was considerably reduced following an identical, but not following a non-identical deviant repetition (Rosburg et al., 2018).

In short, MMN amplitudes underlie habituation if deviants are repeated in directly consecutive trials, especially if the repeated deviants are identical. However, it remains unexplored whether deviant repetition effects interact with the emotional content of the repeated deviant, especially its unpleasantness. Since emotional, especially aversive or painful stimuli, may carry relevant information about potential threats that require a prompt adaptive response, it is plausible to assume that habituation is suppressed for this kind of stimuli compared to emotionally neutral stimuli. For instance, previous studies in the visual domain reported diminished habituation of N1 amplitudes for negative compared to positive or neutral pictures (Carretié et al., 2003), reduced habituation (signaled by the absence of the typical inhibition-of-return effect) for angry compared to happy or neutral faces (Pérez-Dueñas et al., 2014), and decreased adaptation of steady-state visual-evoked potential amplitudes to face identity for fearful compared to happy or neutral faces (Gerlicher et al., 2014). Hence, one could expect a similar suppression of habituation also for sounds that are particularly aversive in nature and potential carriers of threat-related information.

Therefore, we examined potential habituation effects to aversive auditory events during early processing

stages, as reflected in the MMN. Beyond replicating the effects of deviant unpleasantness and deviant repetition on MMN amplitudes, which have been studied separately so far, we aimed to investigate potential interaction effects of these factors. Participants were presented with oddball sequences that contained naturalistic environmental sounds as deviants that were either of aversive or of neutral emotional valence among white noise standards. Deviants occurred as single deviants (following and preceding a standard sound) or as deviant pairs (the same deviant presented in two consecutive trials). This allowed us to measure ERPs depending on the position within deviant pairs, i.e., the effect of habituation, while at the same time controlling for physical stimulus differences. In accordance with earlier findings, we expected (a) enhanced MMN amplitudes in response to aversive as compared to neutral deviants (effect of emotion), and (b) decreased MMN amplitudes in response to deviants in the second as compared to the first position in a deviant pair (habituation). The higher emotional relevance of aversive relative to neutral sounds would render habituation to these sounds evolutionarily maladaptive. Hence, critically, we assumed (c) habituation effects only for neutral, but not or to a lesser extent for aversive deviants. Importantly, we also explored the relationship between participants' valence and arousal ratings of all sound items that occurred as deviants and MMN amplitudes at the item- and participant-level. This allowed us to disentangle whether potential differences between the emotional categories were driven by the stimulus valence or rather by the associated arousal.

2 | METHODS

2.1 | Participants

A total of $N=24$ participants (21 female, three male) took part in the experiment, most of them being undergraduate psychology students at Leipzig University (Germany). They were between 18 and 39 years old ($M=22.46$ years, $SD=4.89$ years), all of them were right-handed, and reported normal or corrected-to-normal vision, normal hearing, and no history of any neurological or psychiatric disorder. Data from all tested participants were included in the analysis. All participants were naïve regarding the purpose of the experiment and gave written informed consent before the testing started. Experimental procedures were performed in accordance with the Declaration of Helsinki, and subjects received either course credits or monetary compensation (8€/hr) for their participation.

2.2 | Stimulation

2.2.1 | Auditory oddball sequences

Auditory stimuli were delivered binaurally via headphones (Sennheiser HD-25-1 II, Sennheiser GmbH & Co. KG, Germany) at approximately 72 dB SPL. The auditory oddball sequences consisted of standard and deviant sounds, each of which had a duration of 300 ms. Stimulus onset asynchrony (SOA) between two consecutive sounds was jittered between 500 ms and 1000 ms (in 10 steps of 50 ms) around a mean SOA of 750 ms. Gaussian white noise served as a standard stimulus, which occurred in 80% of the auditory trials. Deviants were presented in the remaining 20% of the trials and differed with regard to their emotional valence (unpleasantness), such that half of them were neutral and the other half of them were aversive.

Deviants were drawn from a pool of 18 different environmental sounds, which had been selected in a pre-test. In this pre-test, a different group of participants rated the unpleasantness for overall 75 environmental sounds, including natural sounds (such as sounds of birds or water), but also, for instance, cries, metal scraping, or styrofoam squeezing sounds. Based on these ratings and the sounds' characteristics regarding the parameter space of the modeled auditory cortex representation (see Shamma, 2003), we selected a subset of 18 sounds—nine aversive and nine neutral ones. These stimuli were chosen on the criteria of maximizing differences in unpleasantness ratings between the two categories, while minimizing differences regarding the dominant spectral and temporal modulation frequencies in the modeled auditory representations, particularly when sound duration was trimmed to shorter values (i.e., 300 ms). All stimuli were matched in terms of loudness, using the model by Glasberg and Moore for time-varying sounds (Glasberg & Moore, 2002).

Stimuli were arranged in oddball sequences such that the sequence of standards was randomly interspersed with deviants. The latter either occurred as single deviants, i.e., between two standards, or as double deviants, i.e., as a pair of deviants presented in two directly successive trials (see Figure 1). Each of the 18 deviant sounds occurred equally often as a single and as a double deviant. Note that the two deviants within a double deviant pair were always identical, and randomization was restricted such that the second deviant in a pair was always followed by a standard (i.e., no triple deviants occurred). Furthermore, single deviants and double deviants were presented equally often, leading to a 50% probability for either a standard or another (identical) deviant following the first deviant after a row of

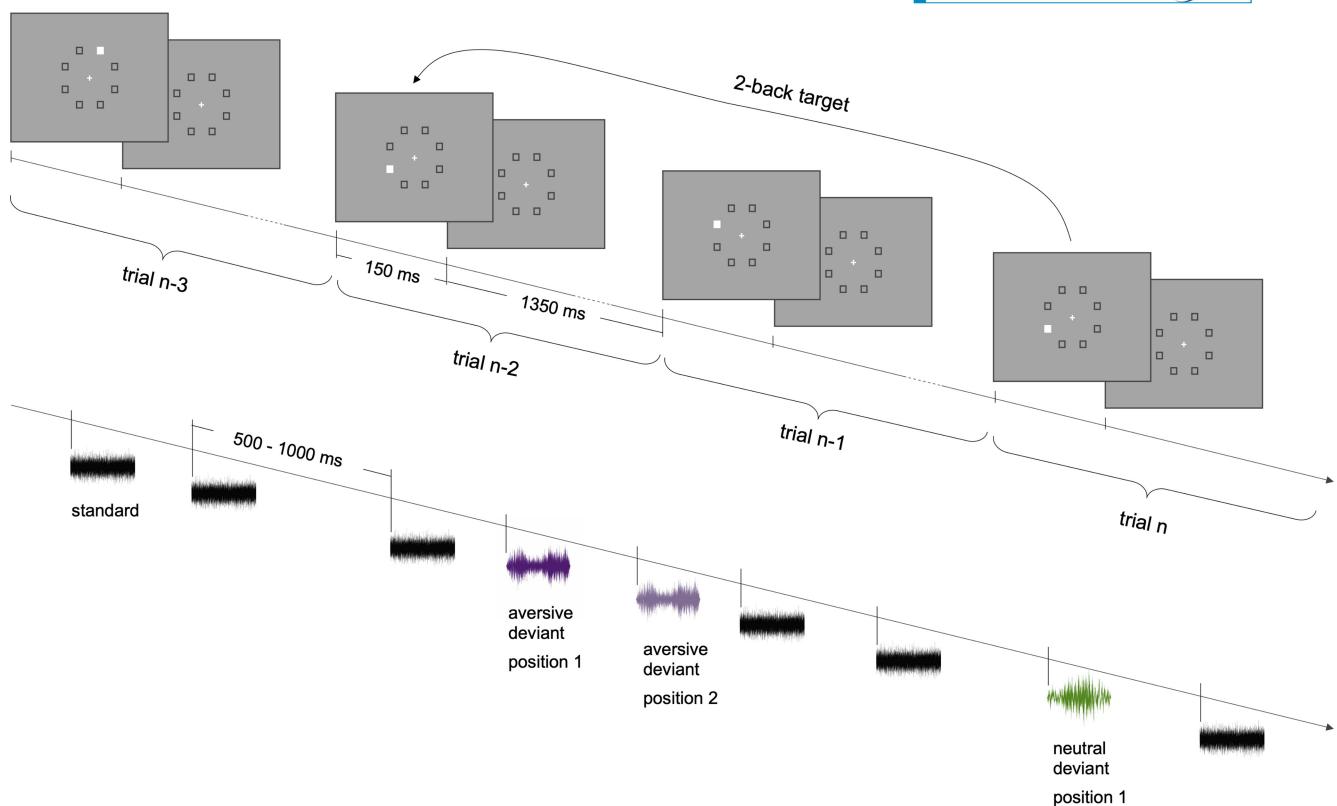


FIGURE 1 Schematic display of concurrent visual and auditory stimulus sequences presented to the participants. In the visual sequence (top of the figure), a white square probe occurred for 150 ms randomly at one of eight possible positions (dark gray frames) every 1500 ms. Concurrently, but unrelated to the visual task, a sequence of 300-ms sounds (bottom of figure) was presented with an SOA jittered between 500 and 1000 ms (in 10 steps). A white noise token (black) was presented in 80% of the auditory trials (standard sound), whereas occasionally either an aversive (purple) or a neutral (green) deviant sound was presented, either in isolation (single deviant) or as a pair (double deviant). The figure shows an example of an aversive double deviant preceded by three standard sounds and a neutral single deviant preceded by two standard sounds. Participants were instructed to focus on a visual 2-back task and to ignore the sounds. (Please note that spatial and temporal proportions were slightly adapted in the figure to improve the visibility).

standards. In short, we manipulated deviants in terms of their unpleasantness (neutral vs. aversive) and position (first vs. second position within a pair), with all types of deviants occurring within the same experimental blocks.

2.2.2 | Visual 2-back distractor task

Visual stimuli were presented on a computer screen (Iiyama HM903DT A, Iiyama North America, Inc.) located outside a shielded cabin behind a window at approximately 80 cm distance from the participants' eyes. The display consisted of a white fixation cross (width and height: 0.37° visual angle) in the center of the screen and eight black-squared frames (width and height: 0.50° visual angle) arranged in a circle around the fixation cross (radius: 2.11° visual angle) at equal distance from each other on a gray background.

During each visual trial, a white square appeared randomly at one of the eight frame positions for 150 ms.

Subjects were asked to fixate the central cross and to compare the position of the white square to the position of the white square two trials before (i.e., 2-back). They were instructed to press a button whenever they detected a 2-back target, that is, when the position of the square in the current trial matched the position of the square two trials before (see Figure 1). The SOA for the appearance of the white squares was set to 1500 ms. 2-back targets occurred randomly in 15% of the visual trials, with the only restriction that each 2-back target was followed by at least two non-target trials.

2.3 | Procedure

During the EEG experiment, participants sat inside a soundproof and electrically shielded cabin. Stimulus presentation was controlled using the Psychophysics Toolbox extension (PTB-3; Brainard, 1997; Kleiner et al., 2007) in Matlab (version R2016a; The MathWorks Inc., USA). Participants' button presses were captured

with a response time box (Suzhou Litong Electronic Co., China). Before the start of the actual experiment, participants had the chance to familiarize themselves with the visual task.

Trials were arranged in 16 blocks. Each block lasted approximately 3.4 min and comprised 270 auditory trials (54 deviant and 216 standard sounds, corresponding to 20% and 80% of the trials, respectively) and 140 visual trials. Thus, over the course of all blocks, each of the 18 deviant sounds was presented 24 times, 8 times as a single deviant and 8 times as a double deviant (with each double deviant corresponding to two sound presentations, respectively). In the first 4 s of each block, only visual stimuli were presented in order to engage participants in the visual distractor task before the auditory sequence started, which they had been instructed to ignore. Auditory and visual trials were neither temporally synchronized nor otherwise systematically related. Detection performance for visual targets was assessed via hit and false alarm rates and reaction times for correct responses to check participants' attention towards the task. We used hit and false alarm rates to compute the sensitivity index d' , defined as the difference of the inverse normal transforms of hit rate minus false alarm rate (corrected using the so-called log-linear transform, see: Hautus & Lee, 2006).

After the EEG experiment, we asked participants to rate the auditory stimuli regarding their valence and arousal on 9-point Likert scales using self-assessment manikins (SAM; Bradley & Lang, 1994). The valence scale ranged from 1 = *very negative* to 9 = *very positive* and the arousal scale ranged from 1 = *not aroused at all* to 9 = *very aroused*. The 18 deviant stimuli and the standard stimulus were presented in random order four times, such that valence ratings were obtained in the first two runs and arousal ratings in the last two runs.

2.4 | EEG data acquisition

During the experiment, EEG data were recorded continuously from 32 Ag/AgCl active electrodes mounted in an elastic cap according to the extended 10–20 system with the following electrode positions: FP1, FP2, AF3, AF4, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, C3, C4, T7, T8, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, PO3, PO4, Oz, O1, and O2. Additionally, signals were recorded from left and right mastoids (M1, M2) and from electrodes placed on the outer canthus of each eye and above and below the right eye (horizontal and vertical electrooculogram). One electrode on the tip of the nose served for later offline referencing. Online high-pass and low-pass filters were applied during recording, attenuating frequencies below 0.16 Hz and above 100 Hz, respectively. Signals (referenced to

the CMS-DRL ground) were amplified with a BioSemi ActiveTwo amplifier (BioSemi B.V., Amsterdam, Netherlands) and digitized with a sampling rate of 512 Hz.

2.5 | EEG data processing and statistical analysis

EEG data were processed offline in Matlab (version R2018b, The MathWorks Inc., USA) with the EEGLAB toolbox (version 15.0.1; Delorme & Makeig, 2004). Statistical analyses were conducted in RStudio (version 4.0.1; R Core Team, 2020). After pre-processing, we determined and compared MMN latencies, and analyzed the MMN amplitudes in two ways: First, to identify the overall effects of deviant unpleasantness and position, we compared amplitudes category-wise between deviants priorly categorized as neutral or aversive that occurred at the first position (following a standard) or at the second position (following a deviant). Second, we evaluated the relationship between stimulus-specific ratings of deviant sounds (in terms of their valence and arousal) and the amplitude of the MMN they elicit when presented at the first or second position.

2.5.1 | Preprocessing

Data were referenced to the electrode located on the tip of the nose and high-pass filtered at 0.2 Hz (transition bandwidth: 0.4 Hz, Kaiser beta: 5.65) and low-pass filtered at 35 Hz (transition bandwidth: 10 Hz, Kaiser beta: 5.65) with Kaiser-windowed sinc finite impulse response filters. We cut the filtered continuous data into epochs that ranged from -100 ms to 500 ms relative to the onset of an auditory stimulus, and baseline-corrected them to the 100 ms interval before sound onset. Epochs corresponding to the first sound that occurred directly after a visual target as well as to standard sounds following a deviant were not included in the analysis. Additionally, any epoch with a peak-to-peak difference exceeding 150 μ V for at least one electrode was discarded from the analysis (on average 13.9% of the epochs).

2.5.2 | Category-wise MMN analysis

After pre-processing, we averaged the remaining epochs within each participant, separately for standards, and for neutral and aversive deviants at the first position after a standard (including both single deviants and deviants at the first position of a deviant pair, because they take functionally indistinguishable roles) and at the second

position within a deviant pair, respectively. We computed difference waveforms by subtracting the standard ERP from the deviant ERP for each of the four deviant types (aversive/neutral \times first/second position). We extracted individual MMN peak latencies—defined as the latency of the minimum within a time window that ranged from 100 ms to 250 ms after stimulus onset—from the difference waveforms at electrode Fz for each deviant type. We statistically compared the extracted MMN latencies between the conditions using a repeated-measures analysis of variance (ANOVA) with the two-level factors Unpleasantness (neutral vs. aversive) and Position (first vs. second position). Mauchly's test of sphericity was used to check whether the assumption of sphericity was fulfilled, and non-sphericity (as indicated by a significant Mauchly's test with $p < .05$) was automatically corrected by applying Greenhouse–Geisser correction.

Subsequently, we computed grand averages across participants from the single-subject difference waveforms for the four deviant types (using the grandaverage plugin for EEGLAB, authored by Andreas Widmann, <https://github.com/widmann/grandaverage>). We determined time windows for the statistical analysis of MMN amplitudes based on the MMN peak latencies in the grand average difference waveforms, again at electrode Fz and within a time window from 100 ms to 250 ms relative to stimulus onset. MMN latencies were shorter for deviants at the first position (neutral: 129 ms; aversive: 121 ms) than for deviants at the second position (neutral: 146 ms; aversive: 156 ms). Therefore, we defined position-specific 40 ms time windows centred around the mean peak latency across deviant unpleasantness levels for the first and second position, respectively. Specifically, MMN time windows ranged from 105 ms to 145 ms for deviants at the first position, and from 131 ms to 171 ms for deviants at the second position. From these time windows, we extracted mean amplitudes at electrodes Fz, M1, and M2 for each subject for each of the four deviant types. To capture the strength of the whole dipole associated with the typical frontocentral negativity and mastoidal positivity, we subtracted the average of the amplitudes at M1 and M2 from the amplitude at Fz. Following this, we submitted the resulting mean MMN amplitudes to a repeated-measures ANOVA with the two-level factor unpleasantness (neutral vs. aversive) and position (first vs. second position).

2.5.3 | Stimulus-wise MMN analysis

We used linear mixed models (built with the lmer function of the lm4 package in RStudio; Bates et al., 2015) to predict MMN amplitudes as a function of the stimulus-specific ratings of the respective deviant sound (in terms

of its valence and arousal) and its position. This approach allowed to model fixed effects of the predictors while controlling for random variance by including random effects (e.g., from random intersubject variance) in the model. Hence, we included participants as a random effect in all our models. Valence and arousal ratings for each stimulus were averaged across the two runs within each participant and mean-centred prior to the linear mixed model analysis. We extracted individual mean MMN amplitudes (again at electrode Fz minus the mean of M1 and M2) for each of the 18 deviant stimuli from position-specific time windows as described above.

In two separate analyses, models were specified analogously for valence and arousal ratings. We computed full models, including the (valence or arousal) rating of a deviant, its position, the interaction of the two factors as fixed effects, and random intercepts for participants to predict MMN amplitudes. We assessed the significance of the fixed effects with *t*-tests using Satterthwaite approximation. Whenever the interaction effect reached significance, we compared the full model (using χ^2 -tests) to a reduced model that considered only (valence or arousal) rating and position as fixed effects, but not their interaction, to test whether the interaction significantly improves the model fit above and beyond the effects of (valence or arousal) rating and position. Furthermore, we compared the full model to reduced models that considered only (valence or arousal) rating or position as the only fixed effect, thus testing whether adding position or (valence or arousal) rating as an additional predictor improved the model significantly.

2.5.4 | Time course of differences between neutral and aversive deviant ERPs

Based on previous studies, our main hypotheses were focussed on the peak of the MMN component, yet processing differences between neutral and aversive deviants seemed to occur already earlier after the sound onset. Therefore, we additionally explored the latency of the earliest processing differences through a non-parametric cluster-based permutation test (Maris & Oostenveld, 2007). The time courses of neutral and aversive deviant ERP (from -100 to 500 ms; averaged across positions 1 and 2) were compared using point-wise independent-samples *t*-tests to identify clusters of amplitude differences between neutral versus aversive deviants (using an alpha of $p < .05$). A permutation approach was used to compare cluster-level *t*-scores against a null hypothesis distribution (using a *tmax* procedure, 1000 permutations, and an alpha level of $p < .05$) in order to determine time intervals of significant differences between neutral and the aversive deviant ERP, specifically in a frontal electrode cluster (around Fz).

3 | RESULTS

3.1 | Visual 2-back distractor task

On average, participants correctly detected 71.9% of the visual targets ($SD=11.8\%$), while they produced very few false alarms (false alarm rate: $M=2.0\%$, $SD=1.1\%$). Individual d' -values ranged from 1.66 to 3.55 around a mean of 2.71 ($SD=0.46$). The mean reaction time for the 2-back target detection was 474 ms ($SD=75$ ms). Together these data suggest that the visual distractor task was challenging, but all participants performed well above chance. These results indicate that their attention was successfully directed towards the task.

3.2 | Ratings of auditory stimuli

In Figure 2 and in Table 1, we show the average valence and arousal ratings (and their standard error of mean or standard deviation) for each of the 18 deviant stimuli as well as for the standard stimulus. Overall, participants tended to rate aversive deviants as more arousing and more negative than neutral deviants. The standard stimulus was located between the two types of deviants on both the valence and the arousal dimension. Notably, participants in the current study rated two of the stimuli that had been categorized as aversive in the previous pilot study (aversive stimulus 1 and stimulus 5) as rather neutral (i.e., valence and arousal ratings were comparable to those of the neutral deviants). Therefore, these two stimuli were excluded from the category-wise ERP analysis.

3.3 | EEG data

Figure 3 shows ERPs (at electrode Fz and averaged bilateral mastoids) evoked by the standard stimulus and by neutral and aversive deviants at the first and second position, as well as the difference waveforms (deviant minus the standard ERP) for each of the four deviant types. The figure also depicts their topography in the MMN time window. All four deviant types elicited a clear MMN as indicated by a frontocentral negativity, along with a polarity reversal at mastoidal electrodes.

3.3.1 | MMN latencies

In the left panel of Figure 4, we depict individual and mean MMN peak latencies at electrode Fz for neutral and aversive deviants at the first and the second position. Across both neutral and aversive deviants, the MMN

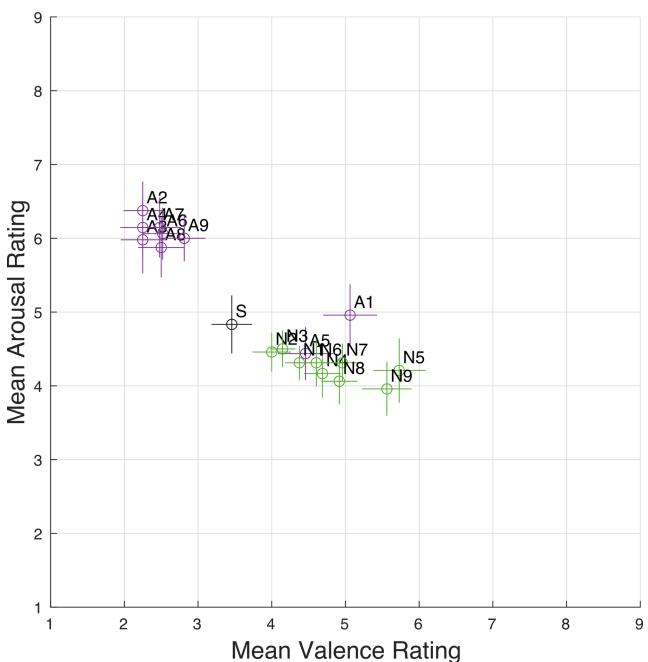


FIGURE 2 Mean valence (x-axis) and arousal (y-axis) ratings for all 18 deviant stimuli and the standard stimulus. Horizontal and vertical bars indicate ± 1 standard error of mean (SEM) on the valence (1 = *very negative* to 9 = *very positive*) and arousal (1 = *not aroused at all* to 9 = *very aroused*) scale, respectively. Aversive sounds are colored in purple (A1 to A9) and neutral sounds in green (N1 to N9) according to the original categorization in a previous pilot experiment. The standard stimulus is colored in black (S). Please note that two of the sounds classified as aversive in the pilot study (marked A1 and A5) were rated as moderately positive and arousing by the current participants. Therefore, responses to these two stimuli were excluded from the category-wise analysis of neutral and aversive deviants.

peaked somewhat earlier in response to deviants at the first position compared to deviants at the second position. The ANOVA with the factors unpleasantness (aversive vs. neutral) and position (1 vs. 2) revealed a significant main effect of position, $F(1, 23)=31.26$, $p<.001$, partial $\eta^2=0.58$, supporting the notion of an overall shorter MMN peak latency for deviants at the first compared to the second position. Conversely, there was neither a main effect of Unpleasantness, $F(1, 23)<0.01$, $p=.945$, partial $\eta^2<0.01$, nor an interaction between the two factors, $F(1, 23)=1.46$, $p=.239$, partial $\eta^2=0.06$, suggesting that the unpleasantness of a deviant does not influence the MMN peak latency.

3.3.2 | MMN amplitudes: Category-wise effects of the deviant unpleasantness and position

In the right panel of Figure 4, we depict individual and mean MMN amplitudes for neutral and aversive

TABLE 1 The mean and standard deviation of the valence and arousal ratings for the auditory stimuli.

	Standard		Aversive deviants		Neutral deviants	
	Valence ratings	Arousal ratings	Valence ratings	Arousal ratings	Valence ratings	Arousal ratings
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Across all stimuli	3.46 (1.50)	4.83 (2.02)	2.95 (1.89)	5.78 (2.15)	4.78 (1.56)	4.25 (1.68)
Individual deviant sounds			5.06 (1.88)	4.96 (2.27)	4.38 (1.16)	4.31 (1.43)
			2.25 (1.44)	6.38 (2.09)	4.00 (1.38)	4.46 (1.38)
			2.25 (1.54)	5.98 (2.26)	4.15 (0.97)	4.50 (1.40)
			2.25 (1.54)	6.15 (2.24)	4.69 (1.32)	4.17 (1.74)
			4.46 (1.96)	4.44 (1.88)	5.73 (1.84)	4.21 (2.20)
			2.52 (1.43)	6.06 (1.95)	4.60 (1.63)	4.31 (1.67)
			2.48 (1.66)	6.15 (2.06)	4.96 (1.49)	4.31 (1.48)
			2.50 (1.57)	5.88 (2.26)	4.92 (1.46)	4.06 (1.77)
			2.81 (1.48)	6.00 (1.68)	5.56 (1.74)	3.96 (1.65)

Note: Average ratings across all nine aversive and neutral deviant stimuli, respectively, as well as for the standard stimulus are shown in the upper row ("across all stimuli"); ratings for the individual sounds are shown in the rows below.

deviants at the first and the second position, respectively. The repeated-measures ANOVA on MMN mean amplitudes with the factors Unpleasantness and Position yielded a significant main effect of Unpleasantness, $F(1, 23)=15.52, p=.001$, partial $\eta^2=0.40$, and a significant main effect of Position, $F(1, 23)=20.18, p<.001$, partial $\eta^2=0.47$, but no significant interaction between the two factors, $F(1, 23)=0.05, p=.830$, partial $\eta^2<0.01$. These results indicate that aversive deviants elicited a larger (i.e., more negative) MMN than neutral deviants at both positions, and deviants at the second position elicited a larger MMN than deviants at the first position, irrespective of the deviant's unpleasantness. Contrary to our expectations, we found no interaction of the two factors on MMN amplitudes.

3.3.3 | MMN amplitudes: Stimulus-wise effects depending on participants' valence and arousal ratings

A substantial proportion of the variance of MMN amplitudes was explained by individual subjects, $ICC(1)=0.16$, $F(23, 840)=7.92, p<.001$, which were therefore modeled as random intercepts. Amplitudes were very homogeneous across participants, $ICC(2)=0.87$.

Valence

The full model, including the valence rating of a deviant, its Position, and the interaction of the two factors as fixed effects as well as random intercepts for participants to predict MMN amplitudes, revealed a significant effect

of Position, $\beta=-1.98, SE=0.23, t(837)=-8.47, p<.001$, but neither a significant effect of Valence rating, $\beta=.14, SE=0.09, t(856)=1.53, p=.127$, nor a significant interaction, $\beta=-.01, SE=0.13, t(837)=-0.08, p=.936$. Subsequent model comparisons showed that the full model, including Valence rating, Position, and their interaction as fixed effects, $AIC=4645.5, R^2=.237$, yielded a significantly better model fit, $\chi^2(2)=69.04, p<.001$, compared to a reduced model with Valence rating as the only fixed effect, $AIC=4710.6, R^2=.172$. Thus, adding position as an additional predictor improved the model substantially. In contrast, including valence rating as an additional predictor did not improve the model, as suggested by the absence of a significant difference in model fit, $\chi^2(2)=3.89, p=.143$, between the full model and a reduced model with Position as the only fixed effect, $AIC=4645.4, R^2=.227$. These results indicate that the position of a deviant reliably predicted the amplitude of the elicited MMN (with larger amplitudes for deviants in the second compared to the first position), whereas no significant additional variance in MMN amplitudes was explained by participants' ratings of a specific deviant sound on the valence dimension (i.e., how positive or negative they perceived the respective sound).

Arousal

The full model, including the arousal rating of a deviant, its position and the interaction of the two factors as fixed effects as well as random intercepts for participants to predict MMN amplitudes, yielded significant effects of both arousal rating, $\beta=-.22, SE=0.09, t(858)=-2.39, p=.017$, and Position, $\beta=-1.98, SE=0.23, t(837)=-8.47, p<.001$, and a

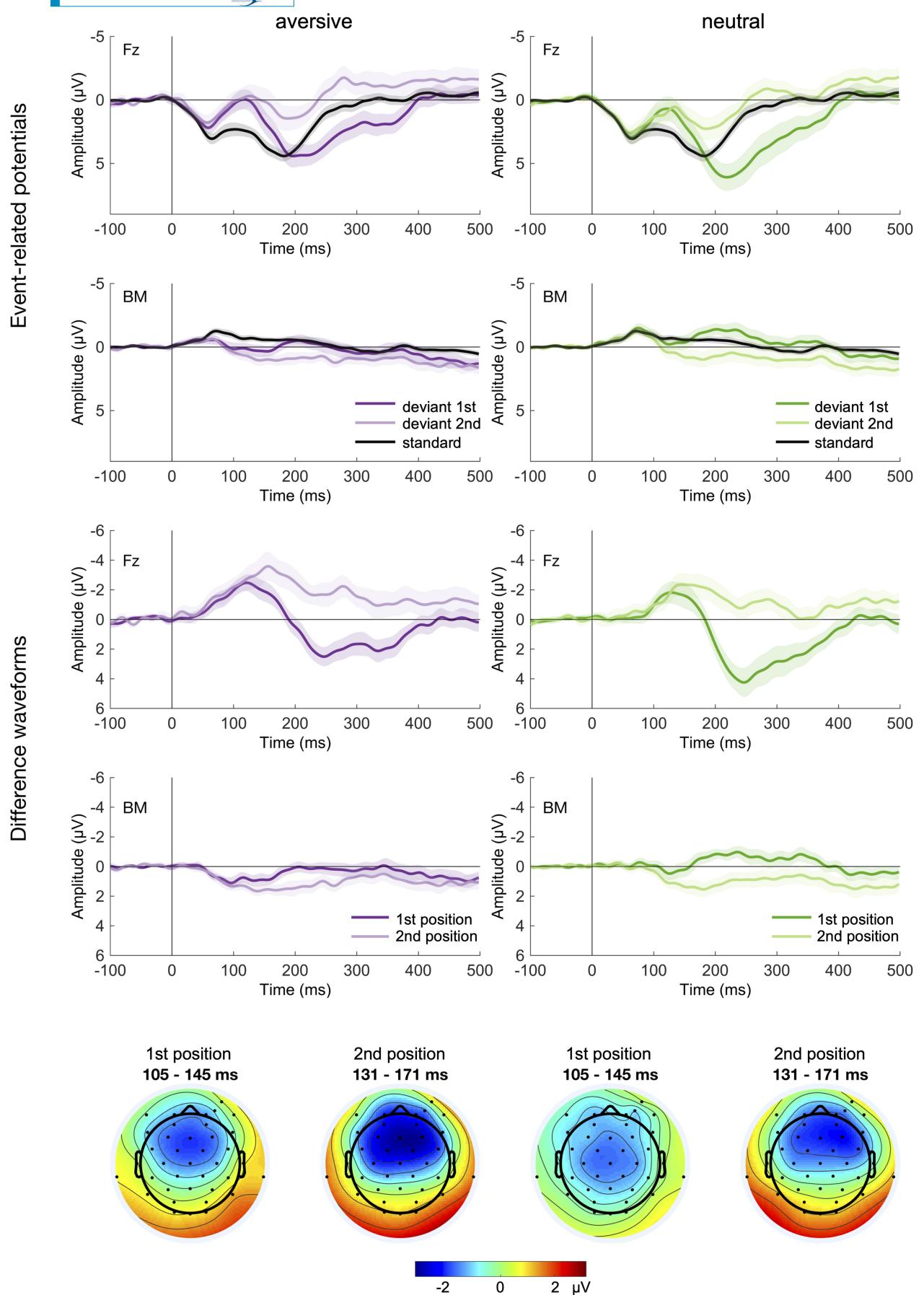


FIGURE 3 Grand-average ERPs (rows 1 and 2 from top) and difference waveforms (rows 3 and 4 from top) for electrode Fz and the average of both mastoids (BM), as well as topographies in the MMN time window (bottom row). We show the data for aversive deviants on the left side, for neutral deviants on the right side.

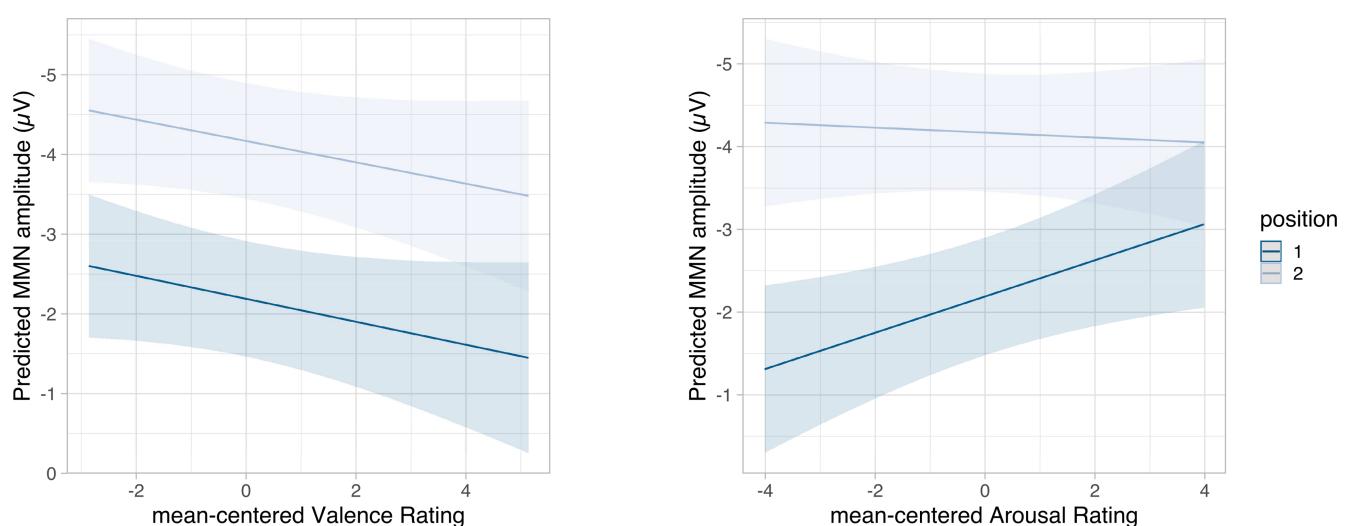
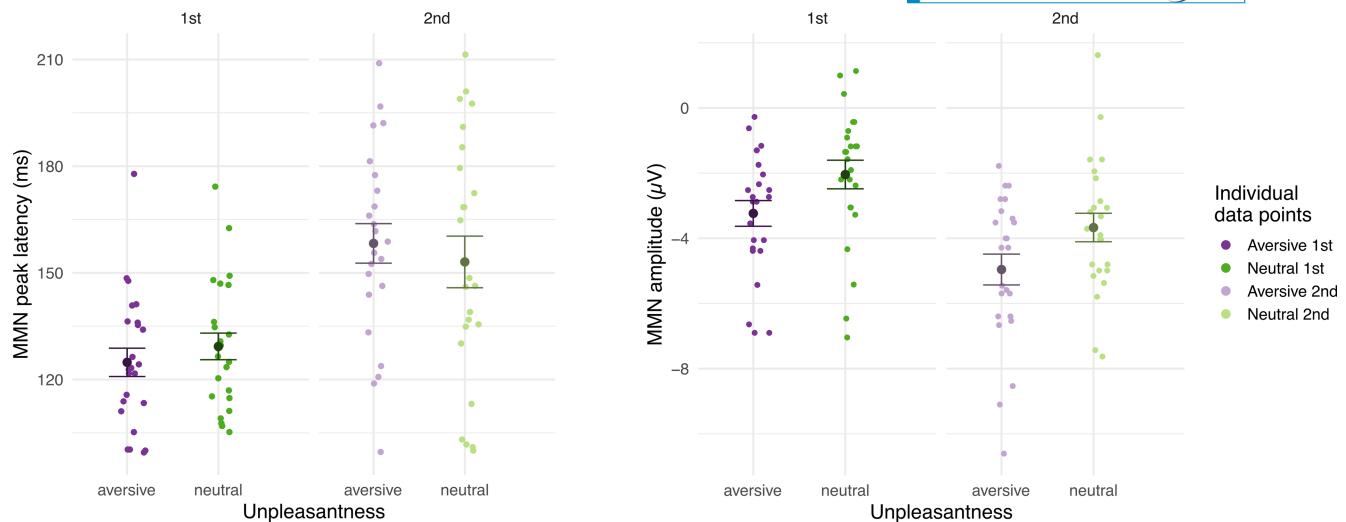


FIGURE 5 Left Panel: Predictions for the MMN mean amplitudes of the 18 deviant sounds presented in the experiment using a linear mixed model including the factors valence rating, position, and valence rating \times position as fixed and participant as random effects. Position reliably predicted MMN amplitudes (with larger amplitudes for deviants at the second than at the first position). Please note that the inclusion of the factor valence rating did not significantly improve the model that included only the factor position. Right panel: Predictions for the MMN mean amplitudes of the 18 deviant sounds presented in the experiment using a linear mixed model including the factors Arousal rating, position, and arousal rating \times position as fixed and participant as random effects. Position reliably predicted MMN mean amplitudes. The factor arousal rating significantly interacted with the position effect. Deviants at the first position were associated with larger MMN amplitudes, when participants rated them as more arousing. For deviants at the second position, arousal ratings did not reliably predict MMN amplitudes.

significant interaction effect, $\beta=.25$, $SE=0.12$, $t(837)=2.04$, $p=.041$. A reduced model that considered only arousal rating and position as fixed effects, but not their interaction, was outperformed by the full model in terms of model fit (full model: $AIC=4643.3$, $R^2=.232$; model without interaction:

$AIC=4645.5$, $R^2=.228$; $\chi^2(2)=4.18$, $p=.041$). Model comparisons further showed that the model fit was significantly better for the full model than for reduced models with either arousal rating, $AIC=4712.3$, $R^2=.163$; $\chi^2(2)=72.96$, $p<.001$, or Position ($AIC=4645.4$, $R^2=.227$; $\chi^2(2)=6.08$, $p=.048$, as

the only fixed effect. This suggests that adding either factor as an additional predictor substantially improved the model. To resolve the significant interaction, we computed separate models including Arousal rating as a fixed effect for the first and for the second position, respectively. As shown in [Figure 5](#), MMN amplitudes were reliably predicted by participants' ratings of a specific deviant sound (i.e., how arousing they perceived the respective sound) at the first, $\beta=-.19$, $SE=0.09$, $t(430)=-2.09$, $p=.037$, but not at the second position, $\beta=.00$, $SE=0.10$, $t(430)=-0.02$, $p=.988$. More arousing deviants elicited a larger MMN when they occurred unexpectedly after a sequence of standard sounds (i.e., at the first position), whereas MMN amplitudes were generally larger and not further modulated by their arousal ratings for repeated deviants (i.e., at the second position).

3.3.4 | Time course of differences between neutral and aversive deviant ERPs

Processing differences between neutral and aversive deviants were seemingly not restricted to the time window of the MMN peak, as neutral deviants elicited a more positive ERP than that of aversive deviants over a wider time interval (see [Figure 6](#)). A cluster-based permutation test revealed that significant ERP differences in an electrode cluster around Fz emerged as early as 50 ms after deviant onset and lasted

until around 320 ms (with a short interval of nonsignificance between 126 and 154 ms after the deviant onset).

4 | DISCUSSION

Using a modified oddball paradigm with single and double deviants we explored how MMN amplitudes are modulated (a) by the aversive nature of deviant stimuli, (b) by (identical) deviant repetition in successive trials, and (c) by potential interactions of those two factors.

Across conditions, all deviants elicited a clear MMN, as reflected by a frontocentrally distributed negativity, accompanied by a polarity reversal at mastoidal electrodes. Crucially, MMN amplitudes were systematically influenced by the deviants' unpleasantness and position. Specifically, our category-wise analysis revealed (a) larger MMN amplitudes for aversive than for neutral deviants (irrespective of their position), and (b) larger MMN amplitudes for the second (identical) than for the first deviant in a row (irrespective of their unpleasantness). The effect of deviant repetition did not differ significantly between both neutral and aversive deviants and might be explained by an overlap with a positive P3a component that only occurred at the first deviant position. Interestingly, using a more fine-grained linear mixed model approach, we observed that deviant position and individual arousal

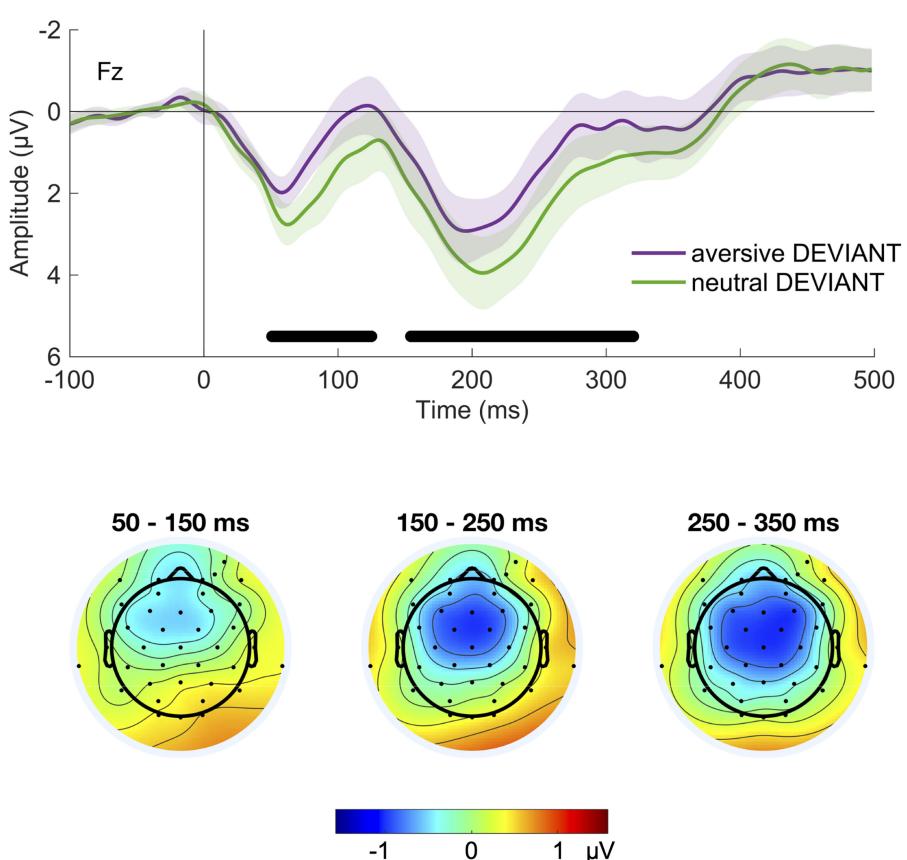


FIGURE 6 Grand-average ERPs for aversive and neutral deviants (averaged across positions) at electrode Fz and topographies of the mean ERP differences (aversive deviant ERP minus the neutral deviant ERP) in three time windows: 50–150 ms, 150–250 ms, and 250–350 ms. The black bar in the upper graph indicates time intervals of significant amplitude differences in the cluster including electrode Fz according to a cluster-based permutation test.

ratings were significant predictors of item-specific MMN amplitudes at the participant-level. In particular, at the first presentation of a deviant, more arousing sounds were associated with larger MMN amplitudes.

4.1 | Increased MMN amplitude for aversive deviants

The finding that MMN amplitudes were enhanced in response to aversive (compared to neutral) deviants is consistent with our hypothesis and closely in line with the results of a previous study (Czigler et al., 2007). It suggests that novel auditory stimuli are discriminated with regard to their unpleasantness at an early stage of processing. In fact, processing differences between neutral and aversive deviants emerged as early as 50 ms after sound onset (i.e., around the time of the beginning slope of the MMN). Aversive deviants may initiate evaluation processes that enable the preparation of adaptive behavioral reactions to potentially relevant or even threatening changes in the environment (Czigler et al., 2007; see also, Schirmer & Escoffier, 2010). Our findings extend those of Czigler et al. (2007) in two critical ways: First, Czigler et al. showed larger MMN amplitudes for aversive than for everyday sounds while participants performed an auditory discrimination task, whereas in the present study participants' attention was directed away from the auditory stimulation towards a visual distractor task. Our findings show distinct central auditory processing for aversive compared to neutral sounds even when attention is directed to another modality, consistent with automatic, pre-attentive processing of aversive sounds. Second, we matched the dominant spectral frequencies and temporal modulation frequencies between neutral and aversive deviants in the stimuli we used. Hence, MMN amplitude differences may be more reliably attributed to the emotional valence and associated arousal of the sounds instead of low-level acoustic properties. Nevertheless, based on the current study we cannot easily distinguish whether the observed differences between neutral and aversive deviants are attributable either to a genuine emotional mismatch response or to the processing of the emotional (aversive) stimulus content per se, which modifies or accompanies auditory deviant processing.

4.2 | Increased MMN amplitude for repeated deviants across unpleasantness levels

Contrary to our hypothesis and in contrast to previous studies (Müller et al., 2005b; Sams et al., 1984), identical deviant repetition did not result in a decrease (i.e., the

expected habituation), but in an increase of MMN amplitudes in response to deviants at the second compared to deviants at the first position. We found this enhancement of MMN amplitudes as a result of deviant repetition for both neutral and aversive sounds. This is also inconsistent with our hypothesis of differential effects depending on the unpleasantness of the repeated deviant. We consider two reasons for this unexpected result.

First, the discrepancy between the present and previous findings might be a consequence of differences in the auditory stimulus material. Whereas earlier studies were conducted with simple sine tones (Müller et al., 2005b; Sams et al., 1984), we employed complex and more ecologically valid environmental sounds that might have led to different repetition effects. In fact, two magnetoencephalography studies provide evidence for enhanced brain responses that emerge around 150 ms after sound onset following the repeated presentation of complex frequency-modulated sounds, but not following the repeated presentation of simpler unmodulated tones (Altmann et al., 2011; Heinemann et al., 2010). Notably, the effect was specific to the first repetition of a frequency-modulated sound and gave way to a second-order repetition suppression during further repetitions (Altmann et al., 2011). It remains to be clarified by future research whether this pattern of initial repetition enhancement and subsequent repetition suppression holds true for complex environmental sounds as used in our study, and whether it may be modulated by the emotional content of the stimuli. Similarly, a recent study compared neural adaptation to complex vocal and non-vocal sounds and found that repetition suppression was delayed, i.e., required more repetitions to occur, for acoustically richer vocal compared to non-vocal sounds, further suggesting that sound complexity may modulate the pattern of repetition effects (Heurteloup et al., 2022).

Second, amplitude differences between conditions in the MMN time window might be confounded with amplitude differences in earlier or later time windows. In our data, both neutral and aversive deviants evoked a positive P3a component from around 200 ms after sound onset for deviants in the first position (cf. Figure 3, best seen in the difference waves at Fz, third row). The P3a component may already begin to emerge at time points overlapping with the MMN time window, thereby partly canceling out the MMN, particularly in conditions with large P3a amplitudes. Given that the P3a seems strongly diminished for deviant repetitions (in line with previous research, e.g., Rosburg et al., 2018), the MMN amplitude differences between first and second deviant position are more likely to reflect an overlap with a larger P3a resulting in a smaller MMN for the first deviant than an increased MMN for the second deviant.

4.3 | Influence of individual arousal, but not valence ratings on MMN amplitudes

Above and beyond the comparison of MMN amplitudes to deviants priorly categorized as neutral or aversive, we evaluated the relationship between individual ratings of sounds and MMN amplitudes when presented as deviants at the first or the second position. Analyses of valence and arousal ratings allowed to dissociate which of these two well-established dimensions to describe and classify emotions (Posner et al., 2005; Russell, 1980) was specifically associated with MMN amplitudes. Interestingly, we did not find any predictive value for MMN amplitude in the individual valence ratings. However, we found that individual arousal ratings reliably predicted MMN amplitudes in interaction with the position of the deviant. MMN amplitudes for deviants were overall larger at the second than at the first position, but more arousing stimuli were associated with enhanced MMN amplitudes for deviants only at the first, but not at the second position. While literature on dissociable effects of valence and arousal in response to auditory stimuli is scarce (see, e.g., Scheumann et al., 2017), more research has been done in the visual domain. For instance, Anders et al. (2004) found that the valence and arousal of emotional pictures were consistently related to activity in distinct brain regions and to distinct peripheral physiological reactions in a functional magnetic resonance imaging study. Vogt et al. (2008), on the other hand, showed that the allocation of (spatial) attention was dependent on the arousal induced by an emotional picture, but not on its valence. Kensinger and Corkin (2004) reported that—compared to neutral words—improved memory capacity for (negative) emotional words relied on distinct neural networks depending on whether the words were arousing: memory capacity for highly arousing, as opposed to non-arousing words, was specifically related to activity in the amygdala irrespective of stimulus valence, which the authors argued to be triggered automatically. If we consider MMN to reflect pre-attentive, automatic stimulus processing, its dependence on or predictability by the arousal associated with a deviant might suggest contributions from brain structures associated with processing emotional or arousing stimuli. For instance, Kumar et al. (2012) showed that aversive sounds activated the amygdala, and that amygdala activity, in turn, modulated activity in the auditory cortex in order to facilitate sensory processing and evaluation of these stimuli. This is also in line with previously reported amygdala activation for emotional compared to neutral voices (Schirmer et al., 2008).

Regarding ERPs, there have been only few attempts to dissociate the effects of valence and arousal. A study by Rozenkrants and Polich (2008) used images from the International Affective Pictures System (IAPS) and found larger

ERP amplitudes (e.g., N2, P3) in response to stimuli associated with high compared to low arousal, whereas stimulus valence had only weak effects on late ERPs. The authors suggested that arousal is the primary determinant of affective oddball processing. Thus, an explanation along the lines of arousal differences that modulate sensory processing appears plausible for our finding of enhanced MMN amplitudes in response to rare, more arousing stimuli.

At first glance, it may seem counterintuitive that our category-wise analysis did not reveal a significant interaction between deviant unpleasantness and position, whereas the stimulus-wise analysis showed a link between MMN amplitudes and participants' arousal ratings that, crucially, depended on the deviant position. However, rather than a discrepancy, the results may reflect a specification of the effect of emotional content on MMN amplitudes: The interaction was driven by one dimension (arousal) of the stimulus' unpleasantness and was specific to individual ratings for the particular sound items (rather than to an *a priori*-defined emotional category). Thus, an interplay between a deviant's unpleasantness and its repetition exists, but apparently acts more subtly at the level of individually perceived arousal than to show in the category-wise analysis.

5 | CONCLUSION

We show evidence for an enhancement of MMN amplitudes in response to aversive compared to neutral environmental sounds presented as deviants. However, given that the identical repetition of deviants resulted in larger MMN amplitudes for the second compared to the first deviant, we did not find signs of habituation as originally expected, but instead observed an unspecific MMN amplitude increase for repeated, complex, and ecologically valid deviants irrespective of their unpleasantness. Critically, we provide evidence that an increase in perceived arousal rather than perceived valence of a deviant sound is associated with a stimulus-specific MMN amplitude enhancement. This strengthens the assumption of a fast emotional evaluation that gates automatic sensory processing with all its evolutionary advantages of allowing adaptive behaviors.

AUTHOR CONTRIBUTIONS

Hanna Ringer: Writing – original draft; writing – review and editing. **Sarah Alicia Rösch:** Data curation; formal analysis; investigation; writing – original draft; writing – review and editing. **Urte Roeber:** Investigation; methodology; resources; software; supervision; writing – review and editing. **Julia Deller:** Data curation; investigation; visualization; writing – review and editing. **Carles Escera:** Conceptualization; methodology; writing – review and editing. **Sabine Grimm:** Conceptualization; data

curation; formal analysis; investigation; methodology; project administration; resources; software; supervision; visualization; writing – review and editing.

ACKNOWLEDGEMENTS

Carles Escera was supported by the ICREA Acadèmia Distinguished Professor award. Sarah Alicia Rösch was funded by a Cusanuswerk scholarship. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available at the Open Science Framework: https://osf.io/2uk78/?view_only=d5dfd1570451445b849586b38405d9a8.

ORCID

Hanna Ringer <https://orcid.org/0000-0001-6115-7878>
 Sarah Alicia Rösch <https://orcid.org/0000-0001-5439-8147>
 Urte Roeber <https://orcid.org/0000-0002-6631-5828>
 Julia Deller <https://orcid.org/0000-0001-5970-4799>
 Carles Escera <https://orcid.org/0000-0001-8203-2951>
 Sabine Grimm <https://orcid.org/0000-0002-9071-5944>

REFERENCES

- Altmann, C. F., Klein, C., Heinemann, L. V., Wibral, M., Gaese, B. H., & Kaiser, J. (2011). Repetition of complex frequency-modulated sweeps enhances neuromagnetic responses in the human auditory cortex. *Hearing Research*, 282(1–2), 216–224. <https://doi.org/10.1016/j.heares.2011.07.008>
- Anders, S., Lotze, M., Erb, M., Grodd, W., & Birbaumer, N. (2004). Brain activity underlying emotional valence and arousal: A response-related fMRI study. *Human Brain Mapping*, 23(4), 200–209. <https://doi.org/10.1002/hbm.20048>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion—the self-assessment mannequin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37(2), 204–215. <https://doi.org/10.1111/1469-8986.3720204>
- Bradley, M. M., & Lang, P. J. (2002). Measuring emotion: Behavior, feeling, physiology. In R. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 242–276). Oxford University Press.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Carminati, M., Fiori-Duharcourt, N., & Isel, F. (2018). Neurophysiological differentiation between preattentive and attentive processing of emotional expressions on French vowels. *Biological Psychology*, 132, 55–63. <https://doi.org/10.1016/j.biopsych.2017.10.013>
- Carretié, L., Hinojosa, J. A., & Mercado, F. (2003). Cerebral patterns of attentional habituation to emotional visual stimuli. *Psychophysiology*, 40(3), 381–388. <https://doi.org/10.1111/1469-8986.00041>
- Chen, C., Sung, J. Y., & Cheng, Y. (2016). Neural dynamics of emotional salience processing in response to voices during the stages of sleep. *Frontiers in Behavioral Neuroscience*, 10, 1–11. <https://doi.org/10.3389/fnbeh.2016.00117>
- Chen, C. Y., Chan, C. W., & Cheng, Y. W. (2018). Test-retest reliability of mismatch negativity (MMN) to emotional voices. *Frontiers in Human Neuroscience*, 12, 453. <https://doi.org/10.3389/fnhum.2018.00453>
- Chen, C. Y., Lee, Y. H., & Cheng, Y. W. (2014). Anterior insular cortex activity to emotional salience of voices in a passive oddball paradigm. *Frontiers in Human Neuroscience*, 8, 743. <https://doi.org/10.3389/fnhum.2014.00743>
- Cox, T. J. (2008). Scraping sounds and disgusting noises. *Applied Acoustics*, 69(12), 1195–1204. <https://doi.org/10.1016/j.apacoust.2007.11.004>
- Czigler, I., Cox, T. J., Gyimesi, K., & Horváth, J. (2007). Event-related potential study to aversive auditory stimuli. *Neuroscience Letters*, 420(3), 251–256. <https://doi.org/10.1016/j.neulet.2007.05.007>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, 10(5), 590–604. <https://doi.org/10.1162/089892998562997>
- Escera, C., & Malmierca, M. S. (2014). The auditory novelty system: An attempt to integrate human and animal research. *Psychophysiology*, 51(2), 111–123. <https://doi.org/10.1111/psyp.12156>
- Gerlicher, A. M. V., van Loon, A. M., Scholte, H. S., Lamme, V. A. F., & van der Leij, A. R. (2014). Emotional facial expressions reduce neural adaptation to face identity. *Social Cognitive and Affective Neuroscience*, 9(5), 610–614. <https://doi.org/10.1093/scan/nst022>
- Glasberg, B. R., & Moore, B. C. J. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5), 331–342.
- Halpern, D. L., Blake, R., & Hillenbrand, J. (1986). Psychoacoustics of a chilling sound. *Perception & Psychophysics*, 39(2), 77–80. <https://doi.org/10.3758/BF03211488>
- Hautus, M. J., & Lee, A. (2006). Estimating sensitivity and bias in a yes/no task. *British Journal of Mathematical and Statistical Psychology*, 59(2), 257–273. <https://doi.org/10.1348/000711005X65753>
- Heinemann, L. V., Rahm, B., Kaiser, J., Gaese, B. H., & Altmann, C. F. (2010). Repetition enhancement for frequency-modulated but not unmodulated sounds: A human MEG study. *PLoS One*, 5(12), e15548. <https://doi.org/10.1371/journal.pone.0015548>
- Heurteloup, C., Merchie, A., Roux, S., Bonnet-Brilhault, F., Escera, C., & Gomot, M. (2022). Neural repetition suppression to vocal and non-vocal sounds. *Cortex*, 148, 1–13. <https://doi.org/10.1016/j.cortex.2021.11.020>
- Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9), 3310–3315. <https://doi.org/10.1073/pnas.0306408101>
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, 36(ECVP Abstract Supplement), 14.

- Kumar, S., Forster, H. M., Bailey, P., & Griffiths, T. D. (2008). Mapping unpleasantness of sounds to their auditory representation. *Journal of the Acoustical Society of America*, 124(6), 3810–3817. <https://doi.org/10.1121/1.3006380>
- Kumar, S., von Kriegstein, K., Friston, K., & Griffiths, T. D. (2012). Features versus feelings: Dissociable representations of the acoustic features and valence of aversive sounds. *Journal of Neuroscience*, 32(41), 14184–14192. <https://doi.org/10.1523/Jneurosci.1759-12.2012>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Müller, D., & Schröger, E. (2007). Temporal grouping affects the automatic processing of deviant sounds. *Biological Psychology*, 74(3), 358–364. <https://doi.org/10.1016/j.biopsych.2006.09.006>
- Müller, D., Widmann, A., & Schröger, E. (2005a). Auditory streaming affects the processing of successive deviant and standard sounds. *Psychophysiology*, 42(6), 668–676. <https://doi.org/10.1111/j.1469-8986.2005.00355.x>
- Müller, D., Widmann, A., & Schröger, E. (2005b). Deviance-repetition effects as a function of stimulus feature, feature value variation, and timing: A mismatch negativity study. *Biological Psychology*, 68(1), 1–14. <https://doi.org/10.1016/j.biopsych.2004.03.018>
- Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313–329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- 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. <https://doi.org/10.1016/j.clinph.2007.04.026>
- Nousak, J. M. K., Deacon, D., Ritter, W., & Vaughan, H. G. (1996). Storage of information in transient auditory memory. *Cognitive Brain Research*, 4(4), 305–317. [https://doi.org/10.1016/S0926-6410\(96\)00068-7](https://doi.org/10.1016/S0926-6410(96)00068-7)
- Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport*, 19(2), 209–213. <https://doi.org/10.1097/WNR.0b013e328f454db>
- Pérez-Dueñas, C., Acosta, A., & Lüpiañez, J. (2014). Reduced habituation to angry faces: Increased attentional capture as to override inhibition of return. *Psychological Research*, 78(2), 196–208. <https://doi.org/10.1007/s00426-013-0493-9>
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3), 715–734. <https://doi.org/10.1017/S0954579405050340>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rosburg, T., Weigl, M., Thiel, R., & Mager, R. (2018). The event-related potential component P3a is diminished by identical deviance repetition, but not by non-identical repetitions. *Experimental Brain Research*, 236(5), 1519–1530. <https://doi.org/10.1007/s00221-018-5237-z>
- Rozenkrants, B., & Polich, J. (2008). Affective ERP processing in a visual oddball task: Arousal, valence, and gender. *Clinical Neurophysiology*, 119(10), 2260–2265. <https://doi.org/10.1016/j.clinph.2008.07.213>
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>
- Sams, M., Alho, K., & Näätänen, R. (1984). Short-term habituation and dishabituation of the mismatch negativity of the ERP. *Psychophysiology*, 21(4), 434–441. <https://doi.org/10.1111/j.1469-8986.1984.tb00223.x>
- Scheumann, M., Hasting, A. S., Zimmermann, E., & Kotz, S. A. (2017). Human novelty response to emotional animal vocalizations: Effects of phylogeny and familiarity. *Frontiers in Behavioral Neuroscience*, 11, 204. <https://doi.org/10.3389/fnbeh.2017.00204>
- Schirmer, A., & Escouffier, N. (2010). Emotional MMN: Anxiety and heart rate correlate with the ERP signature for auditory change detection. *Clinical Neurophysiology*, 121(1), 53–59. <https://doi.org/10.1016/j.clinph.2009.09.029>
- Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport*, 16(6), 635–639. <https://doi.org/10.1097/00001756-200504250-00024>
- Schirmer, A., Escouffier, N., Zysset, S., Koester, D., Striano, T., & Friederici, A. D. (2008). When vocal processing gets emotional: On the role of social orientation in relevance detection by the human amygdala. *NeuroImage*, 40(3), 1402–1410. <https://doi.org/10.1016/j.neuroimage.2008.01.018>
- Schröger, E. (2005). The mismatch negativity as a tool to study auditory processing. *Acta Acustica United with Acustica*, 91(3), 490–501.
- Schweiger Gallo, I., Fernandez-Dols, J. M., Gollwitzer, P. M., & Keil, A. (2017). Grima: A distinct emotion concept? *Frontiers in Psychology*, 8, 131. <https://doi.org/10.3389/fpsyg.2017.00131>
- Shamma, S. (2003). Encoding sound timbre in the auditory system. *IETE Journal of Research*, 49(2–3), 145–156. <https://doi.org/10.1080/03772063.2003.1141633>
- Vogt, J., De Houwer, J., Koster, E. H. W., Van Damme, S., & Crombez, G. (2008). Allocation of spatial attention to emotional stimuli depends upon arousal and not valence. *Emotion*, 8(6), 880–885. <https://doi.org/10.1037/a0013981>

How to cite this article: Ringer, H., Rösch, S. A., Roeber, U., Deller, J., Escera, C., & Grimm, S. (2023). That sounds awful! Does sound unpleasantness modulate the mismatch negativity and its habituation? *Psychophysiology*, 00, e14450. <https://doi.org/10.1111/psyp.14450>