Hearing Research 399 (2021) 107907

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Encoding of deterministic and stochastic auditory rules in the human brain: The mismatch negativity mechanism does not reflect basic probability

Erich Schröger *, Urte Roeber

Institute for Psychology, Leipzig University, Neumarkt 9-19, D-04109, Leipzig, Germany

ARTICLE INFO

Article history:
Received 30 September 2019
Received in revised form
11 January 2020
Accepted 2 February 2020
Available online 7 February 2020

Keywords:
Rule learning
Prediction error
Predictive model
Mismatch negativity (MMN)
Stochastic rule
Deterministic rule

ABSTRACT

Regularities in a sequence of sounds can be automatically encoded in a predictive model by the auditory system. When a sound deviates from the one predicted by the model, a mismatch negativity (MMN) is elicited, which is taken to reflect a prediction error at a particular level of the model hierarchy. Although there are many studies on deterministic regularities, only a few have investigated the brain's ability to encode non-deterministic regularities. We studied a simple stochastic regularity: two tone pitches (standards, each occurring on 45% of trials); this regularity was occasionally violated by another tone pitch (deviant, occurring on 10% of trials). We found MMN when the deviant's pitch was outside those of the standards, but not when it was between them. Importantly, when we alternated the occurrence of the same two standards, making them deterministic, the deviant elicited MMN, even when its pitch was between those of the standards. Thus, although the MMN system is extremely powerful in establishing even quite complex deterministic regularities, it fails with a simple stochastic regularity. We argue that the MMN system does not know basic probability.

1. Introduction

Our brain can automatically encode regularities in the acoustic environment into a model for predicting forthcoming sounds (Winkler et al., 2009; Friston and Kiebel, 2009). The encoding is for simple regularities, such as repetition of a sound (Näätänen et al., 1978), and for more abstract regularities, such as the alternation of two sounds differing in pitch (Nordby et al., 1988). The variety and complexity of regularities that can be learned support the idea of a remarkable "auditory intelligence" (Näätänen et al., 2001). The sequence of sounds forming the regularity does not have to be deterministic. The auditory system tolerates some stochastic variability in the statistical properties of incoming sounds (Garrido et al., 2013; Garrido et al., 2016; Winkler et al., 1990).

The brain's ability to encode stochastic, non-deterministic fluctuations in the distribution of incoming sounds is required in natural situations, because natural soundscapes usually contain some level of randomness (Skerritt-Davis and Elhilali, 2018). Whereas plenty of research is concerned with the complexity of deterministic regularities (for reviews see Näätänen et al., 2001; Paavilainen, 2013), there are only few studies explicitly investigating the brain's ability for encoding non-deterministic regularities. Here, we study the encoding of such regularities into predictive models and systematically compare it with the encoding of deterministic regularities.

The brain's ability to encode regularities into a predictive model has often been discussed (e.g., by Barascud et al., 2016; Lumaca et al., 2019; Wacongne et al., 2012; Winkler et al., 2009) as consistent with predictive-coding theory (e.g., Friston, 2009; Friston and Kiebel, 2009). According to the theory, perception is a result of the interaction between the sensory input and a neural model of the world generating predictions about the sensory input (inferred causes of the input). The difference between the prediction and the actual input is computed as an error signal, which in turn is used to improve the model. The brain hosts a hierarchy of predictive models, with increasing abstraction along the hierarchy. One or several concurrent models can be located at each level along the hierarchy.

The most prominent tool to study predictive models of automatic auditory regularities is the mismatch negativity (MMN),

---

* Corresponding author.
E-mail addresses: schroger@uni-leipzig.de (E. Schröger), urte.roeber@uni-leipzig.de (U. Roeber).

https://doi.org/10.1016/j.heares.2020.107907
0378-5955/© 2020 Elsevier B.V. All rights reserved.

measured with event-related potentials (ERPs) from electroencephalography (EEG) (e.g., Kujala et al., 2007). Garrido et al. (2009) proposed the MMN as reflecting a prediction error signal at a particular level of the hierarchical predictive coding model, a notion attracting wide support (e.g., Wacongne et al., 2012; Winkler et al., 2009; Winkler and Schröger, 2015).

In a typical MMN paradigm, the so-called oddball paradigm, participants are presented with frequent, identical sounds, *standards,* that establish a deterministic regularity, and occasionally, at random, sounds that do not adhere to the regularity, *deviants*. MMN occurs at around 150 ms after the onset of the deviant (for reviews see e.g., Fishman, 2014; Garrido et al., 2009; Lieder et al., 2013; Näätänen, 1992).

Instead, Garrido et al. (2013) randomly sampled the pitch of most of their tones from a Gaussian distribution of given mean and variance. As so-called "probe tones" (Fig. 1, p. 2) they used tones that were either equal to the mean of the distribution or two octaves above it. Garrido et al. found that only the tones outside of the distribution elicited an MMN, with larger amplitude when the variance of the Gaussian distribution was smaller. This important finding reveals the ability of the auditory system to learn non-deterministic, stochastic regularities from an uncertain world and to detect outliers from a given distribution.

We aimed at going one step further by testing whether deviant tones that are not outliers of the distribution spanned by the standards can still be detected as violating the regularity. In other words, we were interested if – in a modification of the oddball paradigm – a predictive model with a bimodal distribution can be established. For simplicity, we used a discrete random variable with three possible outcomes, namely a tone that can be of pitch X, Y, or Z. In the critical "stochastic regularity" condition, the probabilities are .45, .1, and 0.45 respectively. If X, Y, and Z were cards, a gambler would quickly learn to bet on X or Z rather than on Y, because he or she would rightly expect one of these two cards to be more likely to be drawn. How will the auditory system behave? Will it predict the sound sequence to continue with X or Z rather than with Y? Does it matter whether the low-probable tone pitch is in the center of the pitch distribution or at the edge?

More specifically, the sounds in our critical condition were three sine tones with X of 900 Hz, Y of 1000 Hz, and Z of 1100 Hz. In the *stochastic enclosing oddball condition* the highly probable tones X and Z enclosed the rare Y tone in pitch (see Figs. 1 and 4 for illustrations of all experimental conditions). Hence, the pitch of the rare, irregular tone was equal to the mean frequency of the distribution of pitches. We contrasted this critical condition in two experiments against several other conditions. In a *stochastic excluding oddball condition,* the rare, irregular tone was X and the highly probable tones were Y and Z making the irregular tone an outlier of the established distribution of pitches.

We expected to replicate the outlier effect reported by Garrido et al. (2013) for the stochastic excluding oddball condition. In this condition, more or less straightforward predictions about the pitch of the forthcoming sound can be made ("the tone should belong to the range of pitches defined by the established unimodal distribution"). In contrast, for the stochastic enclosing oddball condition, we would expect an MMN only if the auditory system established a

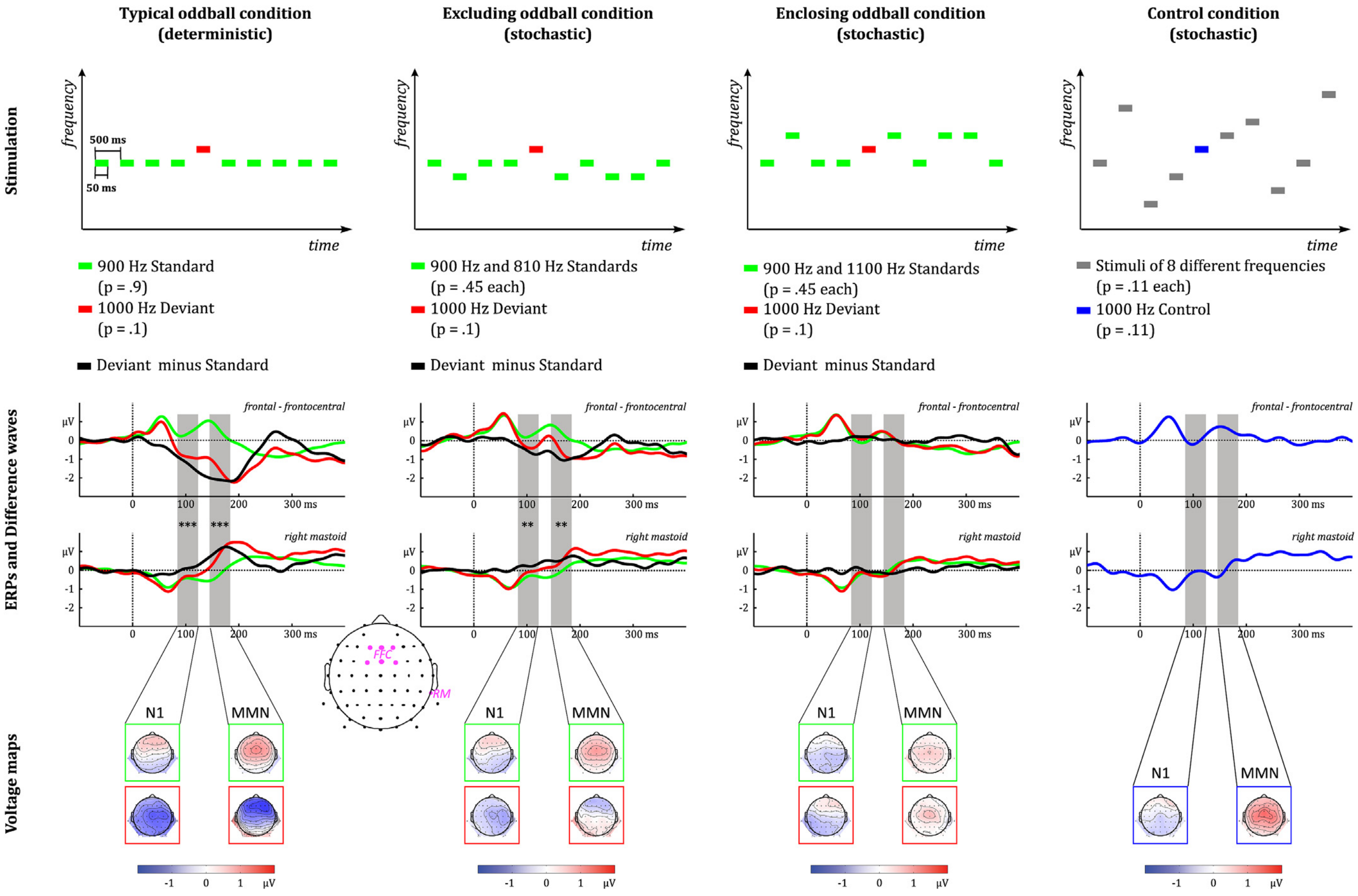

Fig. 1. Top panel: Typical trial sequence for each condition in Experiment 1. Middle panel: Event related potentials (ERPs) and difference waves, from a cluster of six frontal and frontocentral electrodes and from the electrode at the right mastoid (see schematic head between the middle and the bottom panel between the first two conditions). Bottom panel: Voltage maps for standard and deviant stimuli in the N1 and MMN 40-ms time windows.

bimodal distribution of regularity from which an either-or-prediction is made ("the tone should correspond to one of the two highly probable pitches") or if the auditory system established two separate unimodal distributions X and Z from which two concurrent predictions are made (one model says "the next tone is X″ and the other says "the next tone is Z″).

We also ran a typical *deterministic two-tone oddball condition*, in which we combined the rare Y tone with either the frequent X or the frequent Z, as well as a *deterministic alternating enclosing oddball condition*, in which the X and Z tones alternated, rarely replaced by a Y tone. In these deterministic conditions we expected to find MMNs (e.g., Alain et al., 1994; Ritter et al., 1992).

In an additional stochastic enclosing oddball condition, we increased the pitch separation for the X, Y, and Z tones (*stochastic widely enclosing oddball condition*) and compared it with the *stochastic (tightly) enclosing oddball condition* (see above). We used this condition to test whether increasing the pitch separation would restore MMN in a stochastic enclosing condition (if there were no MMN in the stochastic tightly enclosing oddball condition). However, if MMN were restored in this condition, we would expect that this effect is not a genuine MMN, but rather a release from N1 adaptation. In order to separate N1 adaptation from the MMN (reflecting prediction error), we included an *equiprobable control condition* (cf. Näätänen and Alho, 1997; Schröger and Wolff, 1996) in Experiment 1. All these conditions were passive listening conditions, in which participants watched a silenced, subtitled movie or read a book or paper. To test for effects of attention we ran a stochastic enclosing condition in which we asked participants to attend to the sounds and to press a button whenever they heard an irregular tone (*stochastic active-listening tightly enclosing oddball condition).*

2. Method

2.1. Participants

We conducted both experiments in accordance to the Declaration of Helsinki. All participants gave written informed consent for their participation after we explained the nature of the study.

Experiment 1. Twelve participants took part in Experiment 1 for either course credit or money (6 EUR/hour). All reported normal hearing and normal or corrected-to-normal vision. Their ages ranged from 19 to 49 years (median 21 years). All but one of the participants were right-handed. Five participants were male, seven were female.

Experiment 2. Twelve participants took part in Experiment 2 for either course credit or money (6 EUR/hour). All reported normal hearing and normal or corrected-to-normal vision. Their ages ranged from 19 to 29 years (median 23 years). Nine participants were right-handed, three were left-handed. Six participants were male, six were female.

2.2. Apparatus

In both experiments, participants sat comfortably in a sound-attenuated, electrically shielded cabin. Auditory stimuli were delivered binaurally through headphones (70 Ω; HD 25-1, Sennheiser).

In the active-listening condition of Experiment 2, participants responded by pressing one key of a four-key response box.

2.3. Stimuli

In both experiments, stimuli were 50-ms sinusoidal tones (including 5 ms rise and 5 ms fall times). The frequency of the deviant tone was always 1000 Hz. The frequencies of the other tones differed depending on the condition. In Experiment 1, we used frequencies of 656 Hz, 729 Hz, 810 Hz, 900 Hz, 1100 Hz, 1210 Hz, 1331 Hz, and 1464 Hz. In Experiment 2, we used frequencies of 300 Hz, 900 Hz, 1100 Hz, and 1700 Hz.

2.4. Procedure

After attaching the electrodes, we presented the auditory stimuli to the participants while they either watched a silenced, subtitled movie or read a book or paper, except for the active-listening condition in Experiment 2. There were 16 blocks of 540 stimuli in each experiment. Stimulus-onset asynchrony was 500 ms, resulting in a block duration of 4.5 min. We had four conditions in each experiment.

2.4.1. Experiment 1

- *Deterministic oddball condition*: The typical oddball condition contained only one standard frequency, that is, the regularity is deterministic. There were two versions of this condition. In the low-frequency-standards version, 90% of stimuli were 900-Hz tones. In the high-frequency-standards version, 90% of stimuli were 1100-Hz tones. In both versions the standards' frequency excluded the 1000-Hz frequency of the 10% randomly interspersed deviants. We ran four blocks of this condition, two with low-frequency and two with high-frequency standards.
- *Stochastic excluding oddball condition*: The excluding oddball condition contained two standard frequencies. There were two versions of this condition. In the low-frequency-standards version, 45% of stimuli were 900-Hz tones and 45% of stimuli were 810-Hz tones. In the high-frequency-standards version, 45% of stimuli were 1100-Hz tones and 45% of stimuli were 1210-Hz tones. In both versions the standards' frequencies excluded the 1000-Hz frequency of the randomly interspersed deviants. We ran four blocks of this condition, two with low-frequency and two with high-frequency standards.
- *Stochastic enclosing oddball condition*: The enclosing oddball condition contained two standard frequencies. 45% of stimuli were 900 Hz and 45% of stimuli were 1100 Hz. They formed the 90% of standard stimuli that were randomly presented. The standards' frequencies enclosed the 1000-Hz frequency of the 10% randomly interspersed deviants. This is the critical condition of which we ran four blocks.
- *Equiprobable control condition*: The control condition did not contain any standard frequency but stimuli of nine different frequencies (656 Hz, 729 Hz, 810 Hz, 900 Hz, 1000 Hz, 1100 Hz, 1210 Hz, 1331 Hz, and 1464), each occurring randomly with equal probability. That is, 11% of the stimuli were of the same frequency as the deviants in the other conditions (1000 Hz). We ran four blocks of this condition.

We permutated the order of conditions so that each participant was exposed to one block per condition in a different sequence. The same sequence was repeated four times resulting in sixteen blocks. The two versions of the deterministic oddball condition (high or low standard) and the two versions of the stochastic excluding oddball condition (high or low standards) alternated between the mini-sequences of four blocks with the starting version being balanced across participants.

2.4.2. Experiment 2

- *Stochastic tightly enclosing oddball condition*: The tightly enclosing oddball condition contained two standard frequencies that differed by 10% from the deviant frequency: 45% of stimuli were 900 Hz and 45% of stimuli were 1100 Hz. They formed the 90% of standard stimuli that were *randomly* presented. The standards' frequencies enclosed the 1000-Hz frequency of the 10% randomly interspersed deviants. This is a replication of the critical condition in Experiment 1.

- *Stochastic widely enclosing oddball condition*: The widely enclosing oddball condition contained two standard frequencies that differed by 70% from the deviant frequency. 45% of stimuli were 300 Hz and 45% of stimuli were 1700 Hz. They formed the 90% of standard stimuli that were *randomly* presented. The standards' frequencies enclosed the 1000-Hz frequency of the 10% randomly interspersed deviants.
- *Deterministic alternating tightly enclosing oddball condition*: The alternating enclosing oddball condition contained two standard frequencies that differed by 10% from the deviant frequency: 45% of stimuli were 900 Hz and 45% of stimuli were 1100 Hz. They formed the 90% of standard stimuli that were *alternatingly* presented. The standards' frequencies enclosed the 1000-Hz frequency of the 10% randomly interspersed deviants.
- *Stochastic active-listening tightly enclosing oddball condition*: The stimulation in the active-listening tightly enclosing oddball condition was the same as in the tightly enclosing oddball condition, but here participants had to pay attention to the tones and to press a button as soon as they detected a deviant.

Each condition had four blocks that we tested one after the other. We ran the stochastic tightly enclosing oddball condition first and the active-listening conditions last. We counterbalanced the order of the other two conditions among participants.

2.5. Electrophysiological data

We recorded EEG continuously with a BrainAmp system (Brain Products GmbH, Munich) using 66 active Ag/AgCl electrodes (actiCap). Three of them were for monitoring eye movements (Nasion, IO1, and IO2), two were on the left and right mastoids (LM and RM), one was on the tip of the nose. The remaining electrodes were mounted in an elastic cap in positions based on the modified 10–20 system (American Electroencephalographic Society Guidelines in Electroencephalography, Evoked Potentials, and Polysomnography, 1994) (Fig. 1 depicts a schematic head with the electrode positions used). Electrodes were referenced to an electrode on position FCz, and grounded to an electrode on position AFz. The sample rate of EEG and EOGs was 500 Hz.

2.6. Data analysis

Behavioural data. We obtained and analysed behavioural data only for the active-listening tightly enclosing oddball condition in Experiment 2. We calculated d' and reaction times for detected deviants. Because of the high stimulation rate we used the method described by Bendixen and Andersen (2013) to determine the rate of false alarms in order to calculate d'. We defined a deviant being detected when the participant pressed the key between 150 and 1000 ms after the onset of the tone.

Electrophysiological data. To correct for eye movements, we used the method proposed by Schlögl et al. (2007). We re-referenced the EEG data offline to the nose and applied 0.5–35 Hz bandpass filter (Kaiser windowed sinc FIR filter, 1857 points, Kaiser window beta 5.65326) to the data before analysis.

We calculated ERPs by averaging voltages separately for the different kinds of deviants, standards, and equiprobable control stimuli. We used a 600-ms window, time-locked to the onset of the tone, including a baseline from minus 100 to 0 ms. Prior to averaging, we rejected any epoch containing a signal change of more than 100 μV at any EEG electrode by using an automatic peak-to-peak voltage artefact detection method within the epoch window.

Experiment 1. In Experiment 1, we included an average (standard deviation) of 1528 (133) standard trials and 222 (19) deviant trials in the typical oddball condition, of 1471 (154) standard trials and 220 (22) deviant trials in the excluding oddball condition, of 1522 (125) standard trials and 216 (22) deviant trials in the enclosing oddball condition, and of 214 (25) equiprobable control stimuli in the control condition.

For analysis, we calculated difference waves by subtracting the ERPs to standard trials from the ERPs to deviant trials ("traditional" MMN). We determined N1 and MMN time windows by visual inspection of the ERPs and difference waves from the average of six frontal and fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2) forming a frontal-frontocentral electrode cluster (FFC) and from the right mastoid electrode (RM). We decided to show not only the fronto-central region of interest ERPs, but also the right mastoid ERPs for Experiments 1 and 2 for three reasons:

1. In the classical literature, MMN is often presented as a fronto-central negativity and a postero-lateral positivity.
2. In Experiment 2, the mastoid data help to dissociate MMN (which inverts its polarity at mastoid sites) from N2b (which does not invert polarity at mastoid sites).
3. Mastoid data can also help distinguish a deviance-related effect merely at the N1 component (adaptation effect) from a genuine memory-comparison MMN because the N1 should precede MMN on the time scale (with the present stimulus parameters).

In addition to the ERPs, we provide the full topographical information in the voltage maps for the time windows of interest.

For the N1 time window we measured the latency of the first major negative going peak in all ERPs (averaged across participants) at FFC and of the accompanying polarity reversal (positive going peak) at RM, calculated the mean peak latency across all ERPs and electrode sites of all conditions, and defined the N1 time window from 20 ms before to 20 ms after that mean peak latency.

For the MMN time window we measured the latency of the first major negative going peak of the difference waves (averaged across participants) at FFC that was accompanied by a polarity reversal (positive going peak) at RM, of which we also measured the peak latency. We did so only in the conditions that showed these peaks (typical and excluding oddball condition). From these peaks we calculated the mean peak latency and defined the MMN time window from 20 ms before to 20 ms after that mean peak latency.

Following these procedures, the N1 time window ranged from 84 to 124 ms and the MMN time window from 140 to 180 ms. We averaged the amplitudes of the difference waves across the data points of each time window from FFC and from RM and subtracted the mean amplitude at RM from the mean amplitude at FFC in all oddball conditions. We used these mean amplitude differences and *t*-tests to determine N1 effects between standard and deviant amplitudes and the presence of MMN in each oddball condition. We used repeated-measures analysis of variance (ANOVA) to determine effects of condition on the amplitudes of the difference waves.

We performed equivalent calculations for the deviant-minus-equiprobable-control difference waves ("genuine" MMN; Näätänen and Alho, 1997; Schröger and Wolff, 1996).

Experiment 2. In Experiment 2, we included an average (standard deviation) of 1494 (113) standard trials and 213 (16) deviant trials in the tightly enclosing oddball condition, of 1432 (167) standard trials and 203 (23) deviant trials in the widely enclosing oddball condition, of 1422 (224) standard trials and 204 (31) deviant trials in the alternating enclosing oddball condition, and of 1228 (397) standard trials and 116 (71) deviant trials in the active-listening tightly enclosing oddball condition.

For analysis we calculated deviant-minus-standard difference waves. We determined N1, MMN, and N2b time windows by visual inspection of the ERPs and difference waves from the FFC and from the RM.

For the N1 time window and the MMN time window we followed the same procedure as for Experiment 1. That is, for the N1 time window we measured the latency of the first major negative going peak in all ERPs (averaged across participants) at FFC and of the accompanying polarity reversal (positive going peak) at RM, calculated the mean peak latency across all ERPs and electrode sites of all conditions, and defined the N1 time window from 20 ms before to 20 ms after that mean peak latency.

For the MMN time window we measured the latency of the first major negative going peak of the difference waves (averaged across participants) at FFC that was accompanied by a polarity reversal (positive going peak) at RM, of which we also measured the peak latency. We did so only in the condition that showed these peaks (alternating enclosing oddball condition). Because difference waves at RM in the widely enclosing oddball condition and in the active-listening tightly enclosing oddball condition showed a positive going peak in the MMN time window (150–250 ms) we measured their peak latencies as well (even though they were not accompanied by a negative peak at FFC). From these peaks (negativity at FFC in the alternating enclosing oddball condition, and positivities at RM in the alternating enclosing oddball condition, the widely enclosing oddball condition, and the active-listening tightly enclosing oddball condition) we calculated the mean peak latency and defined the MMN time window from 20 ms before to 20 ms after that mean peak latency.

For the N2b time window we measured the latency of the second major negative going peak in the deviant ERPs (averaged across participants) at FFC and RM of the active-listening tightly enclosing oddball condition, calculated the mean peak latency across both electrode sites, and defined the N2b time window from 20 ms before to 20 ms after that mean peak latency. Note, we differentiated between MMN and N2b by looking for negative peaks at FFC as well as at RM, (i.e., showing no polarity reversal).

Following these procedures, the N1 time window ranged from 76 to 116 ms, the MMN time window from 150 to 190 ms, and the N2b time window from 204 to 244 ms. We averaged the amplitudes of the difference waves across the datapoints of each time window from FFC and from RM and, for N1 and MMN we subtracted the mean amplitude at RM from the mean amplitude at FFC in all conditions. We used these mean amplitude differences and *t*-tests to determine N1 effects between standard and deviant amplitudes and the presence of MMN in each oddball condition.

For N2b used the mean amplitudes *t*-tests at FFC and RM to determine effects between standards and deviants in each condition. We used repeated-measures analysis of variance (ANOVA) to determine N1 and MMN effects of condition and N2b effects for condition and electrode location on the amplitudes of the difference waves. When appropriate we used Greenhouse-Geisser corrected degrees of freedom.

Additionally, in the active-listening tightly enclosing oddball condition, we analyzed detected deviant trials and undetected deviant trials separately. Only eight of the twelve participants' data yielded more than fifty trials for both detected and undetected deviant trials. We compared ERPs to detected and undetected deviants and ERPs to standards from these eight participants in the same time windows as above. Average numbers of included trials were 120 (46), 67 (27), and 1317 (287), respectively.

3. Results

3.1. Experiment 1

The top panel of Fig. 1 illustrates a typical trial run for each of the four conditions (from left to right, the three oddball conditions and the equiprobable control). The middle panel shows the ERPs from frontal-frontocentral (FFC) and right-mastoid (RM) electrodes to standard and deviant trials and their difference for the four conditions. The bottom panel shows voltage maps for both stimulus types in each of the four conditions, averaged across five consecutive time windows.

The ERPs to all stimuli show an early positivity (P1) at about 60 ms at FFC that is accompanied by a negativity at RM. P1 amplitudes are about the same for all stimulus types in all conditions. P1 is followed by a negative going peak (N1) at about 90 ms at FFC that is accompanied by a positive going peak at RM for all stimulus types in all conditions. ERPs to standard and deviant stimuli start to differ from about 100 ms after tone onset with deviant stimuli eliciting more negative responses than standard stimuli at FFC. Importantly, these differences occur only in the typical and in the excluding oddball condition but not in the critical enclosing oddball condition.

In Figs. 2 and 3 we collated the difference waves between deviant and standard stimuli of each oddball condition and between deviant stimuli of each oddball condition and the equiprobable control stimuli of the control condition respectively. In both Figures we also show the voltage maps of these ERP differences across the same time windows as above. Both sorts of difference waves show the same pattern of results with no obvious deviant effects in the enclosing oddball condition and a frontal-frontocentral negativity accompanied by a positivity at the mastoids starting at about 100 ms after stimulus onset in the typical and in the excluding oddball condition. These deviant effects were larger in the typical than in the excluding oddball condition.

Statistical comparisons between amplitudes to standard and amplitudes to deviant stimuli by means of difference wave amplitudes (see Table 1) yielded significant FFC minus RM differences in the N1 and MMN time windows in the typical oddball condition and in the excluding oddball condition for the Deviant-minus-Standard differences. As to be expected, in both conditions the N1 effect disappeared when using the Deviant-minus-Control difference, i.e. by subtracting the mean amplitude of the physical identical stimulus of the control condition from the deviant amplitude and thus controlling for adaptation, whereas the MMN effect remained. None of the comparisons was significant for the enclosing oddball condition.

A repeated-measures ANOVA of difference waves amplitudes in the N1 time window showed a significant effect of oddball condition ($F(2,22) = 14.5$, $p < .001$, $\eta^2 = 0.57$) for the Deviant-minus-Standard difference. This is because the Deviant-minus-Standard amplitudes were smaller in the enclosing oddball condition than in the typical and in the excluding oddball conditions ($p < .001$ and $p = .033$, respectively).

In the MMN time window, an equivalent ANOVA also yielded a significant effect of oddball condition ($F(2,22) = 36.34$, $p < .001$, $\eta^2 = .77$) for the Deviant-minus-Standard difference. This is because the Deviant-minus-Standard amplitudes were smaller in the enclosing oddball condition than in the typical and in the excluding oddball conditions ($p < .001$ and $p = .003$, respectively), and Deviant-minus-Standard amplitudes were smaller in the excluding oddball condition than in the typical oddball condition ($p = .002$).

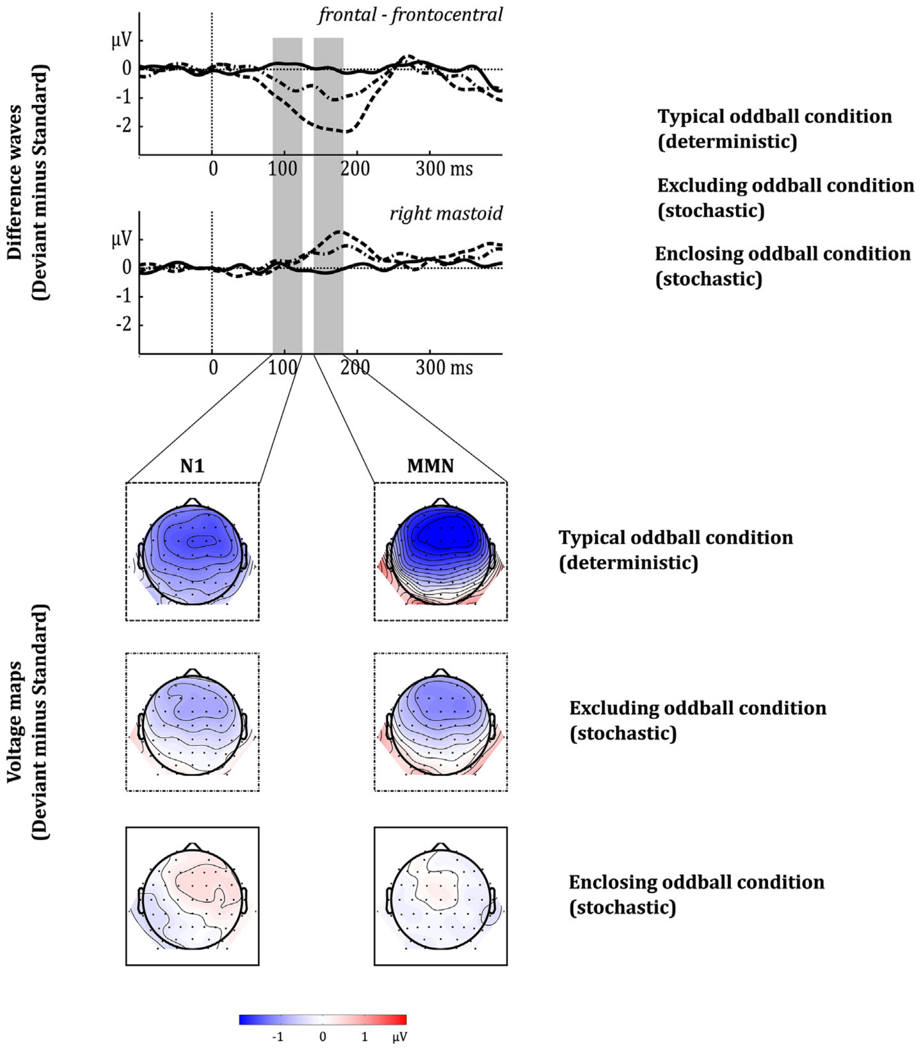

Fig. 2. Difference waves (deviant minus standard) and their voltage maps in the same N1 and MMN 40-ms time windows as in Fig. 1.

3.2. *Experiment 2*

Behavioural data. When participants were asked to detect the tightly enclosed deviants in the active-listening condition in the last four blocks of Experiment 2 their performance varied from one participant not being able to do the task at all to one participant with almost perfect detection. Mean (standard deviation) d′ was 0.78 (1.17) ranging from −1.11 to 3.31. Mean (standard deviation) reaction time of hits was 532 (62) ms.

Electrophysiological data. Fig. 4 illustrates in the top panel a typical trial run and in the middle panel the ERPs from frontal-frontocentral and right-mastoid electrodes to standard and deviant trials and their difference waves in all conditions. The bottom panel in Fig. 4 depicts voltage maps for both stimulus types in each condition averaged across three time windows of interest (N1, MMN, N2b). In Fig. 5 we collated the deviant-minus-standard difference waves of all conditions. We also show the voltage maps of these differences in the same three time windows.

The ERPs to all stimuli show an early positivity (P1) at about 60 ms at the frontal-frontocentral electrodes that is accompanied by a slight negativity at the mastoid electrode. P1 amplitudes are about the same for standards and deviants in each condition. ERPs to standard and deviant stimuli continue to not differ much in the tightly enclosing oddball condition in which P1 is followed by a negativity at about 90 ms (N1) and another positivity (P2) at about 150 ms replicating the results from the enclosing oddball condition in Experiment 1. We also observed N1s and P2s in the other condition. In the widely enclosing oddball condition, deviants elicited larger N1 amplitudes than standards but there is no apparent P2 amplitude difference. In the alternating enclosing oddball condition there is no apparent N1 difference but deviants have smaller P2 amplitudes than standards. In the active-listening tightly enclosing oddball condition, ERPs to deviants and standards did not differ from each other for the first 200 ms. Note, a more specific pattern emerges when we analyzed detected deviants and undetected deviants separately (see below).

In Table 2 we give the results of the statistical comparisons between amplitudes to standard and deviant stimuli in the N1 and MMN time windows for each condition.

In the N2b time window, only the Deviant-minus-Standard amplitudes of the active listening tightly enclosing oddball condition were significantly smaller than zero at FFC ($t(11) = -2.89$, $p = .015$) as well as at RM ($t(11) = -2.46$, $p = .032$).

We performed repeated-measures ANOVAs with the factor condition for the N1 and MMN time windows using the mean amplitudes at RM subtracted from the mean amplitudes at FFC. For N1 we found a significant effect of condition ($F(3,33) = 11.61$, $p < .001$, $\eta^2 = .51$). This is because the Deviant-minus-Standard

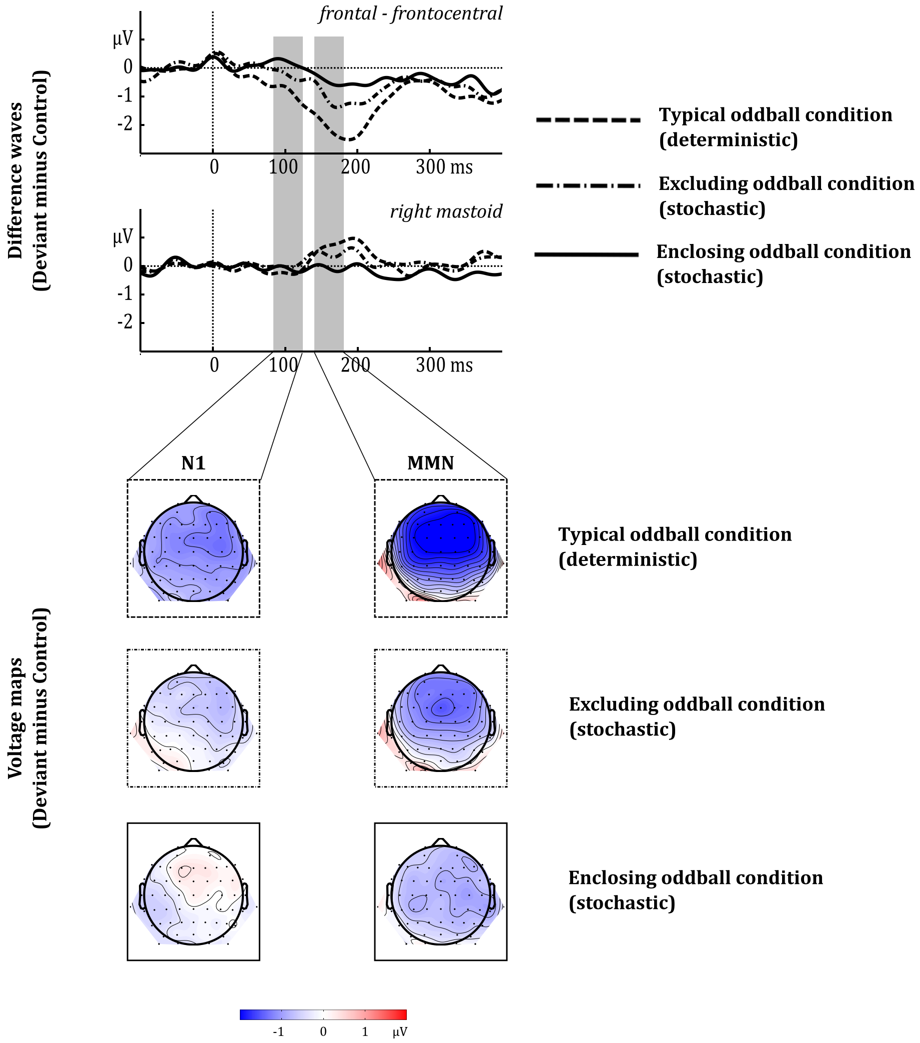

Fig. 3. Difference waves (deviant minus equiprobable control) and their voltage maps in the same N1 and MMN 40-ms time windows as in Fig. 1.

amplitude was significantly larger in the widely enclosing oddball condition as compared to the three other conditions (tightly enclosing oddball condition: $p = .004$; alternating oddball condition: $p = .05$; active-listening tightly enclosing oddball condition: $p = .004$).

For MMN, the effect of condition failed to reach significance after Greenhouse-Geisser correction ($F(3,33) = 3.069$, $p = .061$, $\eta^2 = .22$). Importantly, however, the planned contrast between the tightly enclosing condition and the alternating enclosing condition yielded a significant effect ($t(11) = -2.78$, $p = .028$): amplitudes in the MMN time window were significantly larger in the deterministic alternating enclosing condition than in the stochastic tightly enclosing condition.

We also performed a repeated-measures ANOVA with the factors condition and electrode location for the N2b time window. Here we found significant main effects of both factors (condition: $F(3,33) = 5.22$, $p = .006$, $\eta^2 = .32$; electrode location: $F(1,11) = 5.27$, $p = .042$, $\eta^2 = .432$), but no significant interaction between them ($F(1,11) = 0.24$, $p = .76$, $\eta^2 = .02$). The effect of condition is brought about by the large N2b amplitude in the active-listening condition, the effect of electrode by larger amplitudes at FFC than at RM.

As mentioned above, ERPs differ in the active-listening tightly enclosing oddball condition when a deviant was detected as compared to when the deviant was not detected. In Fig. 6, we show the ERPs and their deviant-minus-standard difference waves averaged across the eight of the twelve participants that had a sufficient number of detected and undetected deviants. Compared with standards detected deviants elicited larger negative deflection as in the 204–244 ms (N2) time window than undetected deviants ($t(7) = -2.12$, $p = .036$).

4. Discussion

In two experiments, we have found evidence that our auditory system does not act like a gambler – at least not in our critical condition in which a low-probable tone pitch was in the center of the pitch distribution and the two surrounding tone pitches followed no other regularity than that they occurred with the same high probability. Successful gamblers know about probability; the MMN mechanism seems not to “know”. The evidence comes from our failure to observe a MMN for the rare deviants in the critical conditions in both experiments.

Importantly, we found MMN with the same stimuli in other experimental conditions:

In the typical deterministic oddball condition of Experiment 1, deviant sounds elicited MMN replicating findings from previous studies (Figs. 1–3; for reviews see Escera and Malmierca, 2014; Näätänen et al., 2005; Sussman et al., 2014). Compared to

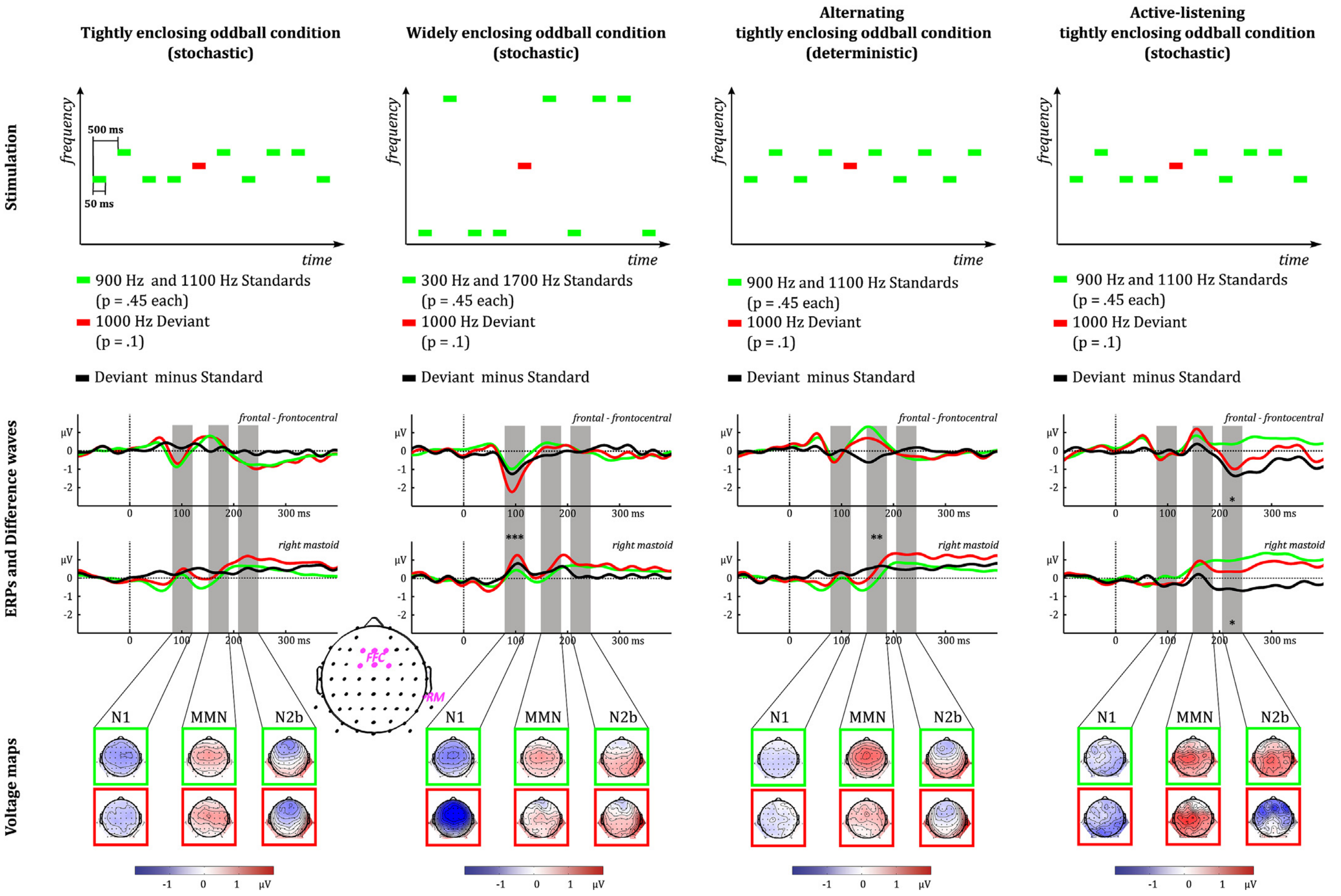

Fig. 4. Top panel: Typical trial sequence for each condition in Experiment 2. Middle panel: Event related potentials (ERPs) and difference waves, from a cluster of six frontal and frontocentral electrodes and from the electrode at the right mastoid (see schematic head between the middle and the bottom panel between the first two conditions). Bottom panel: Voltage maps for standard and deviant stimuli in three 40-ms time windows of interest (N1, MMN, and N2b).

Table 1
t-Values (degrees of freedom = 11) for the comparisons between mean amplitudes to deviants and mean amplitudes to standards by means of difference wave amplitudes of the three oddball conditions (Deviant−Standard) and between mean amplitudes to deviants of the three oddball conditions and the mean amplitude of the physical identical stimulus from the control condition by means of difference wave amplitudes (Control−Standard) from Experiment 1 within the N1 and MMN time windows. All statistics are calculated with the mean amplitudes at RM subtracted from the mean amplitudes at FFC. [two-tailed: $^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$].

	Typical oddball		Excluding oddball		Enclosing oddball	
	Deviant− Standard	Deviant− Control	Deviant− Standard	Deviant− Control	Deviant− Standard	Deviant− Control
N1 (84–124 ms)	-6.14^{***}	-2.18	-3.09^{*}	-0.90	1.18	1.22
MMN (140–180 ms)	-6.34^{***}	-7.34^{***}	-4.02^{**}	-3.56^{**}	0.73	-1.64

standards, deviants elicited an enhanced negativity in the N1 and MMN time range (Fig. 2).

We also found an enhanced negativity in the N1 and MMN time range in the stochastic excluding oddball condition of Experiment 1 replicating the findings of Garrido et al. (2013, 2016) and of Winkler et al. (1990). This emphasizes that the MMN system can encode stochastic regularities into a predictive model with which regularity violations can be detected if the violating sound is at the edge of the pitch distribution. When controlling for adaptation with physical identical stimuli of similar probability in the equiprobable control condition (Schröger and Wolff, 1996) only the differences in the MMN time window (genuine MMN) survives (Fig. 3) in both the typical oddball and the stochastic excluding oddball condition. According to Näätänen et al. (2005), the N1-effect reflects an adaptation mechanism, whereas the MMN reflects a genuine memory-comparison. That we obtained regular MMNs for a deterministic as well as for a stochastic regularity with the present set of tones is also supported by the scalp voltage distributions in the MMN time window. They feature the typical frontocentrally distributed negativity with polarity reversal at occipito-temporal sites below the Sylvian fissure (Figs. 2 and 3).

Our scenario in the stochastic excluding oddball condition uses the simplest stochastic regularity (involving only two pitches to yield a non-deterministic regularity and one pitch serving as deviant). Hence one may argue that any MMN was not because the mechanism detected the stochastic regularity but because of category learning, which has been reported for the speech domain (Goudbeek et al., 2008; Wanrooij and Boersma, 2013). However, if this were the case, one would also expect a MMN for our critical stochastic enclosing oddball condition of Experiment 1 but there was no MMN in this condition. This absence of MMN argues against category learning of two frequent and one infrequent sound.

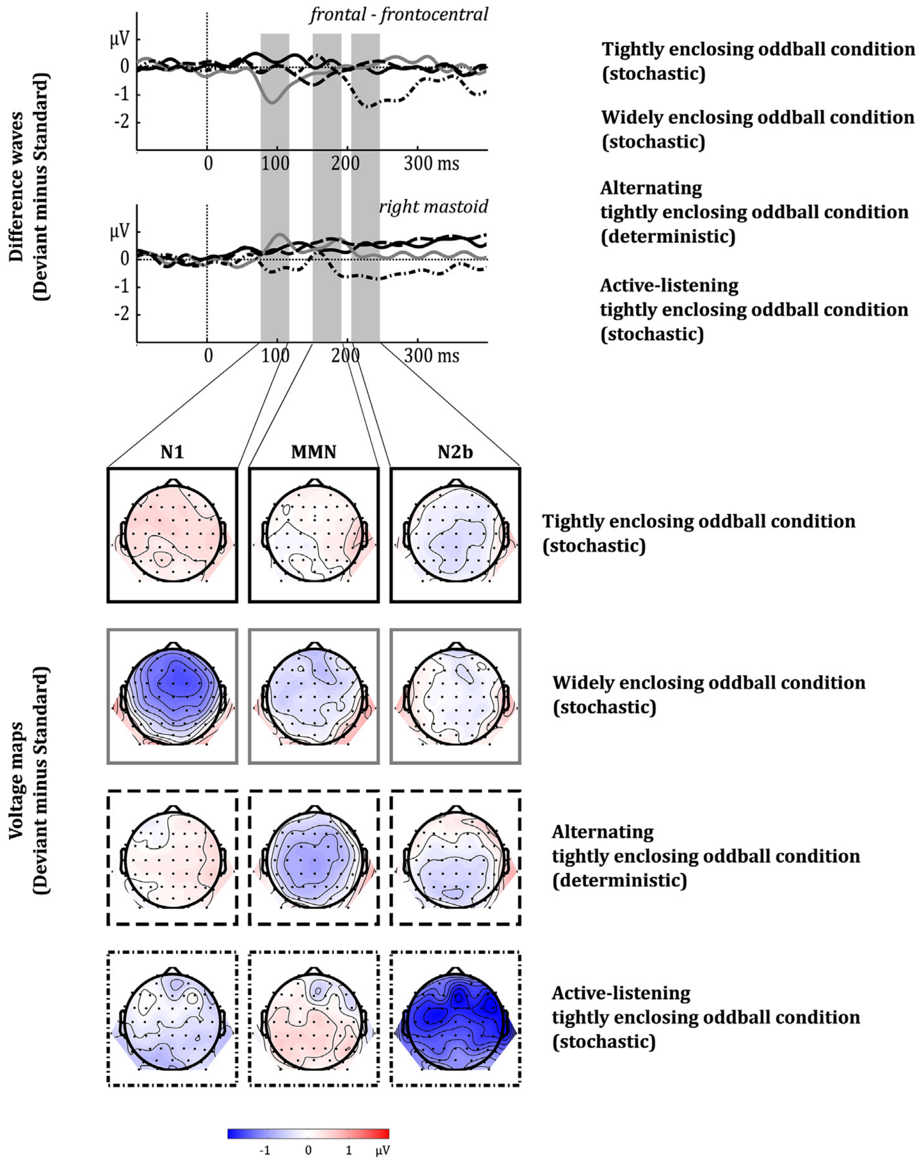

Fig. 5. Difference waves (deviant minus standard) and their voltage maps in the same three 40-ms time windows of interest as in Fig. 4.

Table 2

t-Values (degrees of freedom = 11) for the comparisons between mean amplitudes to deviants and mean amplitudes to standards by means of difference wave amplitudes of the four oddball conditions (Deviant−Standard) and between mean amplitudes to deviants of the three oddball conditions and the mean from Experiment 2 within the N1 and MMN time windows. All statistics are calculated with the mean amplitudes at RM subtracted from the mean amplitudes at FFC. [two-tailed: ***$p < .001$; **$p < .01$; *$p < .05$].

	Tightly enclosing oddball	Widely enclosing oddball	Alternating enclosing oddball	Active-listening tightly enclosing oddball
N1 (76–116 ms)	−0.20	−5.16***	−1.52	1.37
MMN (150–190 ms)	−0.81	−1.87	−3.75**	0.63

The absence of an MMN in the critical condition is consistent with the notion that the stochastic regularity was not encoded by the auditory system. Of course, one cannot rule out the possibility that with longer exposure to this specific soundscape, the stochastic regularity might have eventually been detected and encoded. In fact, long-term training effects have been reported to enhance the MMN amplitude (for a review see Kujala and Näätänen, 2010). If so, this points to a difference in the ease with which the auditory system can encode regularities between the stochastic enclosing oddball condition (difficult) and the stochastic excluding oddball condition (easy).

In predictive coding terms, the results therefore suggest that—with stochastic regularities—the system predicts only that forthcoming sounds will be located within the frequency distribution spanned by the preceding regular sounds. Accordingly, sounds from within the distribution's body – no matter how rare they are – will not be processed as regularity violations, whereas rare sounds outside or at the tails of the distribution will be treated as irregular resulting in a prediction-error response (MMN). For the sake of completeness, it should be mentioned that although the MMN is consistent with predictive-coding theory, it is also consistent a sophisticated adaption framework (e.g., May and

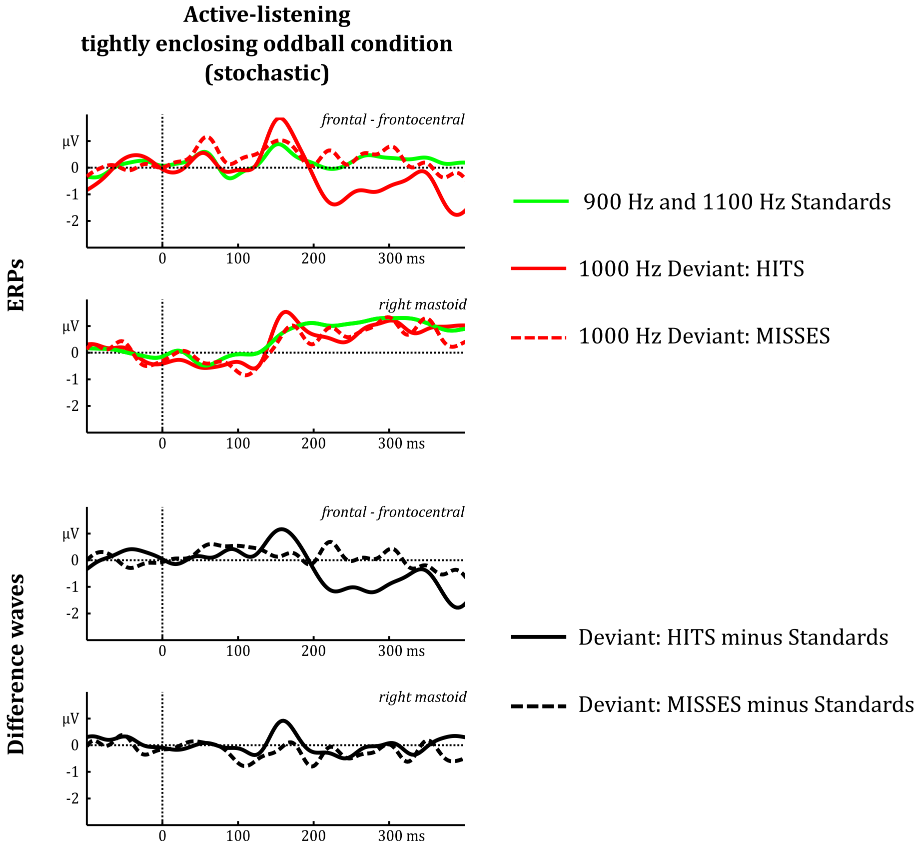

Fig. 6. Event related potentials (ERPs) for standards, detected deviants, (hits) and undetected deviants (misses) and their deviant-minus-standard difference waves, from a cluster of six frontal and frontocentral electrodes and from the electrode at the right mastoid (averaged across eight of the twelve participants).

Tiitinen, 2010) according to which MMN can be explained via fresh-afferent activity of cortical neurons being under non-uniform levels of adaptation. Within this theoretical framework the present results reveal an important constraint on the type of neuronal adaptation underlying the elicitation of MMN.

One would expect a smart regularity-encoding system to be able to extract the information that one of two highly probable sounds is likely to occur in the stochastic regularity conditions and to predict the occurrence of either one of them. We know from other studies that the auditory system can represent several rules in parallel and that it can generate several (Horvath et al., 2001) and even contradictory (Pieszek et al., 2013; Ritter et al., 1999) predictions. Our results show a curious limitation of an otherwise amazingly smart automatic rule-encoding and irregularity-detection system underlying MMN: it fails to make such parallel predictions for the two highly probable sounds in our stochastic regularity condition.

We replicated the results of critical stochastic enclosing oddball condition of Experiment 1 with another set of participants in Experiment 2: we could not find any evidence for a MMN in the data. In Experiment 2 we called this critical condition stochastic tightly enclosing oddball condition. This is because we also ran two other conditions in Experiment 2 to test the hypothesis that MMN did not occur because the pitch difference between tones X, Y, and Z might have been too small: the stochastic widely enclosing oddball condition and the deterministic alternating enclosing oddball condition.

In the deterministic alternating enclosing oddball condition, the same X and Z tones regularly followed each other (alternated) and were rarely replaced by the Y tone. As to be expected from previous reports in the literature (e.g., Alain et al., 1994; Nordby et al., 1988) we found MMN in this condition. Hence the lack of MMN in the stochastic enclosing oddball condition cannot solely be because the pitch differences between the tones were too small. It is important to note that the experimental details were identically between these two conditions, except that the one regularity was deterministic (eliciting MMN), while the other was stochastic (no sign of MMN).

One could still argue that it could be easier for the MMN system to encode the stochastic regularity if the pitch differences were larger. We followed this line of enquiry in the stochastic widely enclosing oddball condition. Indeed, we obtained clear deviance-related effect in this condition seemingly supporting the hypothesis that extraction of stochastic regularity can be facilitated when the sounds become more distinct from each other. However, a closer look at the data reveals that the deviance-related effect was confined to the N1 window. There is a striking similarity between this observation and the deviance-related effect we found in the stochastic excluding oddball condition of Experiment 1. However, in this latter condition we also found a deviance-related effect in the MMN time window that was not present in the stochastic widely enclosing oddball condition of Experiment 2. Because the N1 effect disappeared when contrasting deviants and control tones in Experiment 1, it seems likely that the N1 effect in the stochastic widely enclosing oddball condition of Experiment 2 also reflects an adaptation mechanism rather than a genuine mismatch response or prediction error.

We hasten to add the sound sequence in the stochastic widely enclosing oddball condition did not induce perceptual streaming (due to slow stimulation rate), that is, the perception of a series of X tones and a separate series of Y tones. If this were the case, an MMN should have occurred, because the tone series within each of the two streams would become deterministic (Sussman et al., 1999; Winkler, 2010). Although the present study was designed for determining characteristics of the MMN system, our data show that the N1 system (unlike the MMN system) did know about the stochastic regularity, visible as a huge N1-adaptation effect for the highly probable sound pitches. This points to a theoretically interesting difference in the two deviance-detection systems, which should be addressed in future studies.

We also included an active-listening stochastic tightly enclosing oddball condition in Experiment 2 to test whether attention to the stimulus sequence and an inherent regularity may facilitate the encoding of this (stochastic) regularity. In one way it did, in another it did not. It did, because deviants elicited a negativity in the N2b time window showing that they were detected by the information processing system. However, the voltage map reveals a negative topography even at the mastoid sites Hence this effect should be regarded as an N2b rather than an MMN (Ritter et al., 1992), which is associated with the intentional discrimination of deviants from standards. It did not, because we still failed to observe an unequivocal MMN. So it seems that even with attention the MMN system does not encode stochastic regularities.

We noticed that the behavioral deviance detection performance varied considerably between participants: some participants detected almost all of the deviants; others could not discriminate deviants from standards at all. To test whether the absence of MMN is driven by the participants with bad behavioral performance whose perceptual system might not be able to encode the stochastic regularity we separately analyzed hit and miss trials for each participant (Fig. 6). The analysis revealed a clear N2b (negativity at frontal and mastoid sites) in hit trials but neither component in miss trials, showing the close relationship between conscious discrimination of standard and deviants sounds and the elicitation of N2b. Importantly, however, this analysis did not yield any evidence for MMN, neither in hit nor in miss trials.

Driven by curiosity but with not enough data for a statistical analysis we also looked at the ERPs for the six best behavioral performers (i.e., participants with a *d′* value larger than the median *d′* of 0.95). We show their ERPs for the active-listening stochastic tightly enclosing oddball condition and the deterministic

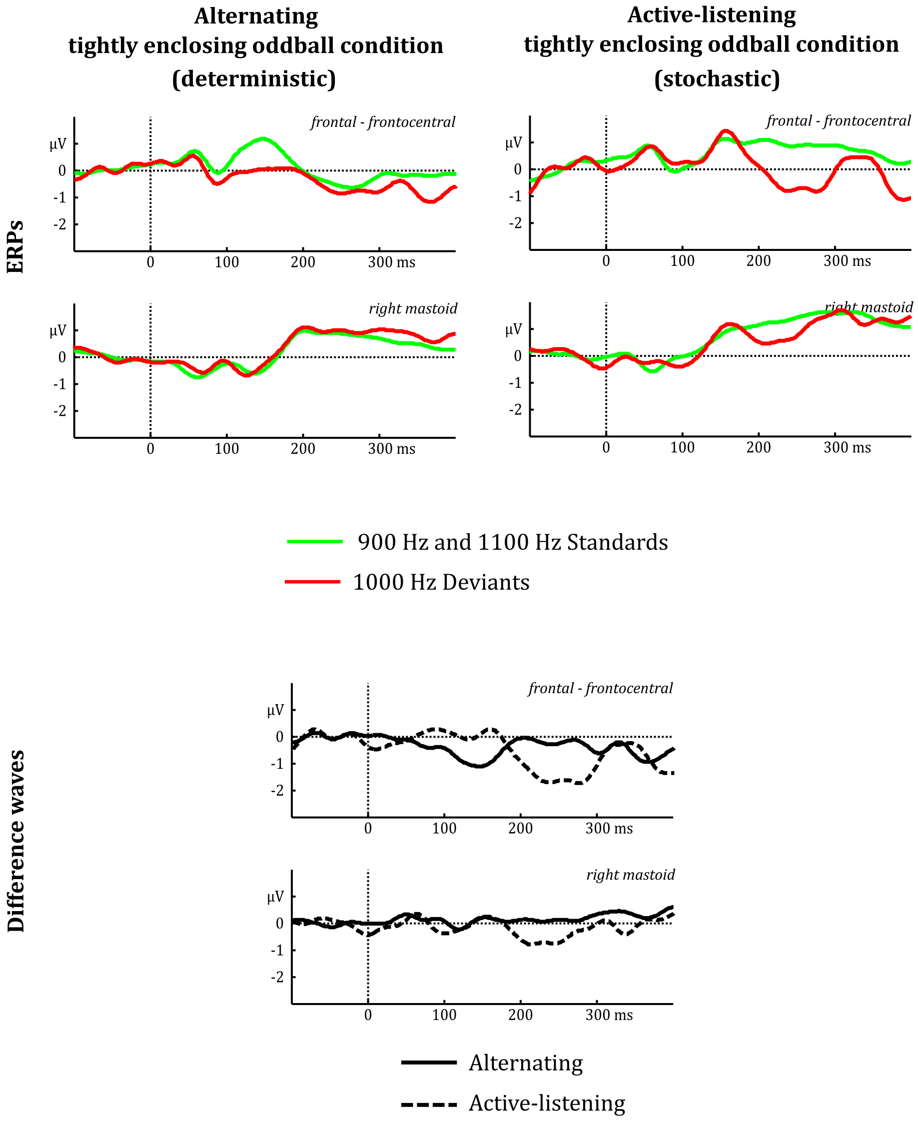

Fig. 7. Event related potentials (ERPs) for standards, detected deviants (hits), and undetected deviants (misses) from a cluster of six frontal and frontocentral electrodes and from the electrode at the right mastoid averaged across the six top performers in the active-listening condition for the deterministic alternating tightly enclosing oddball condition (top left) and for the stochastic active listening tightly enclosing oddball condition (top right). Their deviant-minus-standard difference waves are depicted at the bottom.

alternating enclosing oddball condition in Fig. 7. For the former condition the ERPs yielded an N2b, but show no sign of an MMN. The same participants show a typical MMN in the latter condition. We take this as evidence that a task requiring the encoding of the stochastic regularity does not induce stochastic regularity extraction at the MMN level, but does at a higher level of the model hierarchy, which is reflected by the N2b and in the behavioral performance.

One may ask how the task to discriminate deviants from standards is accomplished by the participants. The participants might simply have encoded the target frequency Y (1000 Hz) and responded to any tone matching this template. Thus, in principle the task could have been accomplished without true deviance detection (in behavioral terms). This kind of behavioral relevance is known to trigger the N2b complex (e.g., Ritter et al., 1992; Sams et al., 1983). In principle, it would have been very easy to control for this explanation by presenting different standard and deviant frequencies in each block or – even better – by controlling for it via a roving standard paradigm (e.g., Baldeweg et al., 2004). However, we chose to have the very same standard and deviant tone frequencies X, Y and Z in the critical stochastic and deterministic conditions during passive and active listening for two reasons: (1) in order to being able to compare the results from passive and from active listening and (2), more importantly, in order to test whether attention would regain MMN with the same stimulus parameters, that did not yield MMN during passive listening.

The participants might have also accomplished the task by learning the two highly probable sounds X and Z, matching the incoming sound with these templates, and not to respond when there is a match, but to do so when there is no match. Again, this kind of behavioral relevance is known to trigger the N2b complex (e.g., Ritter et al., 1992; Sams et al., 1983). A third scenario is that participants behaved like a gambler and predicted (at a more cognitive/conceptual level than the one tapped by the MMN) that either sound X or sound Z will occur next; if this prediction is wrong, N2b is elicited. In this case, the N2b could be regarded as a prediction-error signal. Although we regard this third alternative of how the behavioral task has been accomplished as the most likely, we did not design the present study to decide between these alternatives but to test for the elicitation of MMN by violations of stochastic and non-stochastic regularities.

Strictly speaking our findings imply that it may not be the encoding of stochastic regularities that underlies MMN in the existing stochastic MMN studies (e.g., Garrido et al., 2013) but rather the learning of a category. This category is defined by the mean and the variance of a unimodal distribution of the preceding tones. The unimodal distribution of the preceding tones establishes a stable transitional probability of sounds belonging to this high-probable category despite the fact that they are physically different from each other. This means that perceptual categories can be derived from stochastic regularities (in unimodal distributions), but that the MMN system needs serial regularities in order to be able to make predictions about forthcoming sounds.

So, why can a gambler quickly learn to bet on the highly probable events X or Z rather than on the low probable event Y, and why could (at least some of) our participants well discriminate deviant Y from standard X and Z tones as indicated by behavioral performance and N2b brain responses? It might be that categories based on bimodal distributions can be encoded more readily on a cognitive than on a sensory level.

We argue that inter-sound relationships matter (Mittag et al., 2016; Winkler and Schröger, 2015). The MMN system highly depends on transitional probabilities between adjacent and even non-adjacent (Bendixen et al., 2012) sounds rather than solely on the probability of the sounds. It fits nicely with recent evidence provided by Skerritt-Davis and Elhilali (2018) who showed that the elicitation of MMN can be better explained with higher-order statistics of temporal dependence rather than with lower-order statistics (mean and variance) alone, which may result in adaptation effects. This hypothesis is also supported by recent successful approaches modelling the brain as a near-optimal inference device that analyses and applies the transition probabilities between the stimuli it receives (Maheu et al., 2019; Meyniel et al., 2016). The importance of inter-sound relationships has been considered in many pattern and streaming MMN studies (e.g., Alain et al., 1994; Bendixen et al., 2010; Brattico et al., 2002; Sussman et al., 1998; Winkler and Schröger, 1995) as well as in a computational model of perceptual sound organizations related to MMN (Mill et al., 2011, 2013).

5. Summary and conclusions

We replicated the results from previous studies obtaining MMN with deterministic regularities. We also replicated previous findings of MMN with stochastic regularities (Garrido et al., 2013, 2016; Winkler et al., 1990). Noteworthy, however, we found evidence that the ability to encode a stochastic regularity into a predictive model has limitations. We did not find MMN indicating prediction error when a deviant sound was located in the midst of the distribution spanned by the stochastic regularity. Neither an increase in the physical separation of the pitch difference between the sounds (facilitating their perceptual separation) nor the allocation of attention to the regularity by making deviants task-relevant restored the MMN.

We propose that the failure of the MMN mechanism to use what every gambler knows is a principle constraint of the regularity encoding at the MMN level that might be explained by higher-order statistics of inter-sound relationships. We conclude that the MMN system – unlike other regularity encoding systems indicated by N1 and N2b - does not apply basic probability.

Author contributions

Conceived the experimental idea: ES. Designed the experiments: ES, UR. Performed the experiments: UR. Analyzed the data: UR. Wrote the paper: ES, UR.

CRediT authorship contribution statement

Erich Schröger: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing - original draft, Writing - review & editing. **Urte Roeber:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgements

We thank ten University of Leipzig third-year students for their help and enthusiastic participation in the laboratory project for Experiment 1 in 2008. We are grateful to Jenny Kokinous and Inka Bretschneider who assisted with data collection. We thank Robert P. O'Shea for helpful comments on previous versions of the paper. The experiment was realized using Cogent 2000 developed by the Cogent 2000 team at the FIL and the ICN and Cogent Graphics developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience. The data were analyzed using the MatLab toolboxes EEGLab, and ERPLab. Financial support came from the German Research Foundation (DFG, Reinhart-Koselleck Project SCH 375/20-1 awarded to Erich Schröger).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heares.2020.107907.

References

Alain, C., Woods, D.L., Ogawa, K.H., 1994. Brain indices of automatic pattern processing. Neuroreport 6 (1), 140–144. https://doi.org/10.1097/00001756-199412300-00036.

American Electroencephalographic Society Guidelines in Electroencephalography, Evoked Potentials, and Polysomnography, 1994. J. Clin. Neurophysiol. 11 (1), 1–147.

Baldeweg, T., Klugman, A., Gruzelier, J., Hirsch, S.R., 2004. Mismatch negativity potentials and cognitive impairment in schizophrenia. Schizophr. Res. 69 (2–3), 203–217. https://doi.org/10.1016/j.schres.2003.09.009.

Barascud, N., Pearce, M.T., Griffiths, T.D., Friston, K.J., Chait, M., 2016. Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. Proc. Natl. Acad. Sci. U.S.A. 113 (5), E616–E625. https://doi.org/10.1073/pnas.1508523113.

Bendixen, A., Andersen, S.K., 2013. Measuring target detection performance in paradigms with high event rates. Clin. Neurophysiol.: Off. J. Int. Fed. Clin. Neurophysiol. 124 (5), 928–940. https://doi.org/10.1016/j.clinph.2012.11.012.

Bendixen, A., Denham, S.L., Gyimesi, K., Winkler, I., 2010. Regular patterns stabilize auditory streams. J. Acoust. Soc. Am. 128 (6), 3658–3666. https://doi.org/10.1121/1.3500695.

Bendixen, A., Schröger, E., Ritter, W., Winkler, I., 2012. Regularity extraction from non-adjacent sounds. Front. Psychol. 3, 143. https://doi.org/10.3389/fpsyg.2012.00143.

Brattico, E., Winkler, I., Näätänen, R., Paavilainen, P., Tervaniemi, M., 2002. Simultaneous storage of two complex temporal sound patterns in auditory sensory memory. Neuroreport 13 (14), 1747–1751. https://doi.org/10.1097/00001756-200210070-00011.

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.

Fishman, Y.I., 2014. The mechanisms and meaning of the mismatch negativity. Brain Topogr. 27 (4), 500–526. https://doi.org/10.1007/s10548-013-0337-3.

Friston, K., 2009. The free-energy principle: a rough guide to the brain? Trends Cognit. Sci. 13 (7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Phil. Trans. Roy. Soc. Lond. B Biol. Sci. 364 (1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300.

Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin. Neurophysiol. 120 (3), 453–463. https://doi.org/10.1016/j.clinph.2008.11.029.

Garrido, M.I., Sahani, M., Dolan, R.J., 2013. Outlier responses reflect sensitivity to statistical structure in the human brain. PLoS Comput. Biol. 9 (3) https://doi.org/10.1371/journal.pcbi.1002999.

Garrido, M.I., Teng, C.L.J., Taylor, J.A., Rowe, E.G., Mattingley, J.B., 2016. Surprise responses in the human brain demonstrate statistical learning under high concurrent cognitive demand. Npj Sci. Learn. 1, 16006. https://doi.org/10.1038/npjscilearn.2016.6.

Goudbeek, M., Cutler, A., Smits, R., 2008. Supervised and unsupervised learning of multidimensionally varying non-native speech categories. Speech Commun. 50 (2), 109–125. https://doi.org/10.1016/j.specom.2007.07.003.

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

Kujala, T., Näätänen, R., 2010. The adaptive brain: a neurophysiological perspective. Prog. Neurobiol. 91 (1), 55–67. https://doi.org/10.1016/j.pneurobio.2010.01.006.

Kujala, T., Tervaniemi, M., Schröger, E., 2007. The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. Biol. Psychol. 74 (1), 1–19. https://doi.org/10.1016/j.biopsycho.2006.06.001.

Lieder, F., Stephan, K.E., Daunizeau, J., Garrido, M.I., Friston, K.J., 2013. A neurocomputational model of the mismatch negativity. PLoS Comput. Biol. 9 (11), e1003288 https://doi.org/10.1371/journal.pcbi.1003288.

Lumaca, M., Trusbak Haumann, N., Brattico, E., Grube, M., Vuust, P., 2019. Weighting of neural prediction error by rhythmic complexity: a predictive coding account using mismatch negativity. Eur. J. Neurosci. 49 (12), 1597–1609. https://doi.org/10.1111/ejn.14329.

Maheu, M., Dehaene, S., Meyniel, F., 2019. Brain signatures of a multiscale process of sequence learning in humans. ELife 8. https://doi.org/10.7554/eLife.41541.

May, P.J.C., Tiitinen, H., 2010. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47 (1), 66–122. https://doi.org/10.1111/j.1469-8986.2009.00856.x.

Meyniel, F., Maheu, M., Dehaene, S., 2016. Human inferences about sequences: a minimal transition probability model. PLoS Comput. Biol. 12 (12), e1005260 https://doi.org/10.1371/journal.pcbi.1005260.

Mill, R.W., Bőhm, T., Bendixen, A., Winkler, I., Denham, S.L., 2011. CHAINS: competition and cooperation between fragmentary event predictors in a model of auditory scene analysis. In: 2011 45th Annual Conference on Information Sciences and Systems, pp. 1–6. https://doi.org/10.1109/CISS.2011.5766095.

Mill, R.W., Bőhm, T.M., Bendixen, A., Winkler, I., Denham, S.L., 2013. Modelling the emergence and dynamics of perceptual organisation in auditory streaming. PLoS Comput. Biol. 9 (3), e1002925 https://doi.org/10.1371/journal.pcbi.1002925.

Mittag, M., Takegata, R., Winkler, I., 2016. Transitional probabilities are prioritized over stimulus/pattern probabilities in auditory deviance detection: memory basis for predictive sound processing. J. Neurosci. 36 (37), 9572. https://doi.org/10.1523/JNEUROSCI.1041-16.2016.

Näätänen, R., 1992. Attention and Brain Function. Lawrence Erlbaum Associates, Inc.

Näätänen, R., Alho, K., 1997. Mismatch negativity—the measure for central sound representation accuracy. Audiol. Neuro. Otol. 2 (5), 341–353. https://doi.org/10.1159/000259255.

Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. 42 (4), 313–329.

Näätänen, R., Jacobsen, T., Winkler, I., 2005. Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. Psychophysiology 42 (1), 25–32. https://doi.org/10.1111/j.1469-8986.2005.00256.x.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., Winkler, I., 2001. "Primitive intelligence" in the auditory cortex. Trends Neurosci. 24 (5), 283–288. https://doi.org/10.1016/s0166-2236(00)01790-2.

Nordby, H., Roth, W.T., Pfefferbaum, A., 1988. Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals. Psychophysiology 25 (3), 262–268.

Paavilainen, P., 2013. The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: a review. Int. J. Psychophysiol.: Off. J. Int. Org. Psychophysiol. 88 (2), 109–123. https://doi.org/10.1016/j.ijpsycho.2013.03.015.

Pieszek, M., Widmann, A., Gruber, T., Schröger, E., 2013. The human brain maintains contradictory and redundant auditory sensory predictions. PloS One 8 (1), e53634. https://doi.org/10.1371/journal.pone.0053634.

Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., Näätänen, R., 1992. Event-related potentials to repetition and change of auditory stimuli. Electroencephalogr. Clin. Neurophysiol. 83 (5), 306–321. https://doi.org/10.1016/0013-4694(92)90090-5.

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

Sams, M., Alho, K., Näätänen, R., 1983. Sequential effects on the ERP in discriminating two stimuli. Biol. Psychol. 17 (1), 41–58. https://doi.org/10.1016/0301-0511(83)90065-0.

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., Pfurtscheller, G., 2007. A fully automated correction method of EOG artifacts in EEG recordings. Clin. Neurophysiol. 118 (1), 98–104. https://doi.org/10.1016/j.clinph.2006.09.003.

Schröger, E., Wolff, C., 1996. Mismatch response of the human brain to changes in sound location. Neuroreport 7 (18), 3005–3008.

Skerritt-Davis, B., Elhilali, M., 2018. Detecting change in stochastic sound sequences. PLoS Comput. Biol. 14 (5), e1006162 https://doi.org/10.1371/journal.pcbi.1006162.

Sussman, E., Ritter, W., Vaughan, H.G., 1998. Attention affects the organization of auditory input associated with the mismatch negativity system. Brain Res. 789 (1), 130–138. https://doi.org/10.1016/s0006-8993(97)01443-1.

Sussman, E., Ritter, W., Vaughan, H.G., 1999. An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology 36 (1), 22–34. https://doi.org/10.1017/s0048577299971056.

Sussman, E.S., Chen, S., Sussman-Fort, J., Dinces, E., 2014. The five myths of MMN: redefining how to use MMN in basic and clinical research. Brain Topogr. 27 (4), 553–564. https://doi.org/10.1007/s10548-013-0326-6.

Wacongne, C., Changeux, J.-P., Dehaene, S., 2012. A neuronal model of predictive coding accounting for the mismatch negativity. J. Neurosci. 32 (11), 3665–3678. https://doi.org/10.1523/JNEUROSCI.5003-11.2012.

Wanrooij, K., Boersma, P., 2013. Distributional training of speech sounds can be done with continuous distributions. J. Acoust. Soc. Am. 133 (5), EL398–EL404. https://doi.org/10.1121/1.4798618.

Winkler, I., 2010. In search for auditory object representations. In: Czigler, I., Winkler, I. (Eds.), Unconscious Memory Representations in Perception: Processes and Mechanisms in the Brain. John Benjamins Publishing, pp. 71–106.

Winkler, I., Denham, S.L., Nelken, I., 2009. Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cognit. Sci. 13 (12), 532–540. https://doi.org/10.1016/j.tics.2009.09.003.

Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., 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 (2), 228–235. https://doi.org/10.1111/j.1469-8986.1990.tb00374.x.

Winkler, I., Schröger, E., 1995. Neural representation for the temporal structure of sound patterns. Neuroreport 6 (4), 690–694. https://doi.org/10.1097/00001756-199503000-00026.

Winkler, I., Schröger, E., 2015. Auditory perceptual objects as generative models: setting the stage for communication by sound. Brain Lang. 148, 1–22. https://doi.org/10.1016/j.bandl.2015.05.003.