

# Action Intention-based and Stimulus Regularity-based Predictions: Same or Different?

Betina Korka<sup>1</sup>, Erich Schröger<sup>1</sup>, and Andreas Widmann<sup>1,2</sup>

## Abstract

■ We act on the environment to produce desired effects, but we also adapt to the environmental demands by learning what to expect next, based on experience: How do action-based predictions and sensory predictions relate to each other? We explore this by implementing a self-generation oddball paradigm, where participants performed random sequences of left and right button presses to produce frequent standard and rare deviant tones. By manipulating the action–tone association as well as the likelihood of a button press over the other one, we compare ERP effects evoked by the intention to produce a specific tone, tone regularity, and both intention and regularity. We show that the N1b and Tb components of the N1 response are modulated by violations of tone regularity only. However, violations of action intention as well as of regularity elicit MMN responses, which

occur similarly in all three conditions. Regardless of whether the predictions at sensory levels were based on either intention, regularity, or both, the tone deviance was further and equally well detected at hierarchically higher processing level, as reflected in similar P3a effects between conditions. We did not observe additive prediction errors when intention and regularity were violated concurrently, suggesting the two integrate despite presumably having independent generators. Even though they are often discussed as individual prediction sources in the literature, this study represents to our knowledge the first to directly compare them. Finally, these results show how, in the context of action, our brain can easily switch between top–down intention-based expectations and bottom–up regularity cues to efficiently predict future events. ■

## INTRODUCTION

Predicting forthcoming sensory input allows us to act efficiently in the environment. According to the predictive coding theory, the human brain is a probability calculator, constantly preoccupied with predicting future events (Knill & Pouget, 2004). ERPs can be interpreted as a measure of prediction error, where attenuated sensory ERPs indicate smaller prediction errors and thus better predictions (Friston, 2005). In the auditory domain, several types of prediction signatures, along with their paradigms of investigation, are discussed in the literature (Schröger, Marzecová, & SanMiguel, 2015; Bendixen, SanMiguel, & Schröger, 2012; Hughes, Desantis, & Waszak, 2012). Two prominent lines focus on action-based predictions investigated in self-generation paradigms and sensory predictions investigated in variants of the oddball paradigm.

We act to produce desired outcomes in the environment; according to the ideomotor theory, performing an action results in an association between the action itself and its sensory consequences, and once the association has been learned, action selection is determined postdictively based on its corresponding perceptual consequences (Elsner & Hommel, 2001; Prinz, 1997). Thus, our

own actions represent top–down information sources used to generate predictions. In this context, self-generated tones (most often via button presses) are commonly found to elicit attenuated N1 and often P2 ERP responses in comparison to externally generated but otherwise identical tones (Horváth, 2015). The frontocentral N1b and the temporal Tb peak of the T-complex represent N1 subcomponents that have associated with sensory-specific predictions, in contrast to the “unspecific” N1 observable with large ISIs (SanMiguel, Todd, & Schröger, 2013; Hari, Kaila, Katila, Tuomisto, & Varpula, 1982). Even though the N1–P2 are often discussed together in self-generation studies, it has been suggested that the P2 reflects different processes compared with the N1 (Crowley & Colrain, 2004), which are rather related to processing of complex tone features (Shahin, Roberts, Pantev, Trainor, & Ross, 2005).

It has been proposed that the intention for action (rather than the action itself) is the crucial prediction input—specifically, Timm and colleagues showed that only voluntary button presses, in contrast to involuntary ones induced by TMS, lead N1–P2 attenuation, in comparison to externally generated tones (Timm, SanMiguel, Keil, Schröger, & Schönwiesner, 2014). However, the self-external comparison is problematic because it confounds several processes (for a detailed description, see Hughes et al., 2012). Hughes and colleagues addressed this problem

<sup>1</sup>University of Leipzig, <sup>2</sup>Leibniz Institute for Neurobiology, Magdeburg, Germany

and further compared self-generated tones, which were either congruent or incongruent with hand-specific learned associations and showed that the congruent relative to the incongruent tones, indeed lead to N1 attenuation (Hughes, Desantis, & Waszak, 2013).

Not only do we act to change the environment, but we also adapt to it by learning the relationship between certain events—we thus know what outcome to expect next based on probabilities. In this context, tone regularity represents a source of bottom-up predictive information, automatically extracted from preceding sensory input. It was indeed shown that if the precise identity of the self-generated tones is not stable between trials, the prediction effect is reduced or even abolished (SanMiguel, Saupe, & Schröger, 2013; SanMiguel, Widmann, Bendixen, Trujillo-Baretto, & Schröger, 2013; Bäß, Jacobsen, & Schröger, 2008). Although less often discussed in the context of action, variants of the oddball paradigm investigating sensory predictions as differences between regularity-violating and regularity-conforming tones are vast and well established (for a comprehensive review, see Näätänen, Paavilainen, Rinne, & Alho, 2007). The main component of interest in oddball paradigms is the MMN, which represents the difference between the rare deviant and frequent standard tones, peaking in between 100 and 250 msec after the occurrence of the deviancy (see Garrido, Kilner, Stephan, & Friston, 2009, for a predictive coding interpretation of the MMN).

Although “motor and sensory predictions may constitute different sources for a single mechanism” (Lange, 2013), they are barely ever integrated into a common perspective. This study takes a step forward by considering intention- and regularity-based predictions as distinct prediction sources for action-related predictions, in the context of self-generated sounds. Specifically, we first wanted to determine effects of the violation of the predictions for a particular sound that were either based on the action intention (participant intentionally generated a particular sound as effect of a particular action) or on the presentation regularity (one of the two sounds was presented more frequently than the other, without a reliable action–effect coupling). Second, we wanted to see whether bottom-up regularity-based and top-down intention-based predictions have additive effects in case of concurrent violations. To this end, we used a self-generation paradigm where participants pressed buttons to produce frequent standard and infrequent deviant tones, while we manipulated the action–tone association as well as the likelihood of a button press over the other one.

The N1 and MMN components have been considered as the same brain response (Jääskeläinen et al., 2004), but also fundamentally different (Näätänen, Jacobsen, & Winkler, 2005). Moreover, the incongruity response, overlapping in latency and morphology to the MMN, is yet another prediction error elicited by incongruent audio-visual pairs (Pieszek, Widmann, Gruber, & Schröger, 2013; Widmann, Kujala, Tervaniemi, Kujala, & Schröger, 2004). A

clear distinction between these responses is thus difficult. Here, we use temporal PCA for the ERP analysis, which, in contrast to visual inspection, reliably identifies the constituent wave components, given the complex nature of an ERP wave (Dien, 2012). We shall therefore focus the analysis on the obligatory components (negative and positive, as identified in the data), rather than on the difference wave.

Finally, according to the stages of auditory distraction, if the prediction errors reflected at the N1 and MMN levels reach a strong enough threshold, a second processing stage involving an involuntary attention switch occurs, which is reflected in the P3a component (Horváth, Winkler, & Bendixen, 2008; Waszak & Herwig, 2007; Escera, Alho, Winkler, & Näätänen, 1998; Schröger, 1997). This effect presumably represents the ERP signature of an orienting response following motivationally significant stimuli, such as expectancy-violating deviant tones (Nieuwenhuis, De Geus, & Aston-Jones, 2011). We test the hypothesis that tone regularity as well as action intention lead to sensory prediction effects, as reflected by significant N1/MMN differences between self-generated deviants and standards. If the two prediction sources are additive, this should lead to larger deviant–standard differences in case of concurrent violations. Note that by looking at the differences between standard (predicted) and deviant (unpredicted) tones, we do not measure prediction directly, but we probe the existence of predictions via the effects of prediction violations. However, for reasons of simplifying, we should regard the effects obtained by prediction violation as a measure of prediction. Additionally, we expected to find P3a enhancement for the deviant as compared with standard tones, provided the prediction errors elicited at the earlier processing level reach a strong enough threshold.

## METHODS

### Participants

Data were collected from 24 participants (10 men, mean age = 23.5 years, age range = 18–32 years), all of whom gave written informed consent for the study participation. All participants reported normal hearing and normal-to-corrected vision, and none of them had any history of neurological conditions. All participants were right-handed, except one left-handed man. None was taking any prescribed drugs. The ethics committee of the University of Leipzig, in agreement with the Declaration of Helsinki, approved the study procedure (code of approval: 465/17-ek). Participants received compensation of either €8/hr or course credits.

### Stimuli and Apparatus

For the whole experiment duration, participants were seated in a comfortable office chair in an electrically shielded, double-walled sound booth (Industrial Acoustics

Company). Stimuli were complex sine wave sounds with a fundamental frequency of 352 Hz (the low tone) and 440 Hz (the high tone), including the second and third harmonic attenuated by  $-3$  and  $-6$  dB, respectively, with the duration of 100 msec, including 5-msec rise-and-fall times, and presented binaurally over a pair of headphones (Sennheiser HD 25) at an intensity level of 76 dB SPL. The two keys participants pressed had dimensions of  $6 \times 6$  cm and were placed on a desk, in front of them. Visual feedback indicated how many times the left versus right keys need to be pressed, as well as the time interval between two consecutive key presses (presented in white numbers on a black screen). This was provided on a 19-in. CRT monitor (G90fB, ViewSonic, resolution  $1024 \times 768$  pixels, refresh rate of 100 Hz), which was placed at a comfortable watching distance in front of the participant ( $\sim 60$  cm). Stimuli were created and presented via the Psychtoolbox 3 (Kleiner et al., 2007), in combination with GNU Octave Version 4.0.0 (Eaton, Bateman, Hauberg, & Wehbring, 2016), running on Linux OS.

### Task

Participants pressed the left and right keys using their left and right index fingers to generate tones, according to the condition-specific instructions. Figure 1 displays a possible condition-specific representation of key press–tone associations. For the Regularity condition (Figure 1A), participants were instructed to press both keys with 50–50% chances to generate a low tone (which was presented in 80% of the cases). Rarely, a high tone was presented instead (20% of the cases). For the Intention condition (Figure 1B), participants' instructions were to press the left key in 50% of the cases to generate a low tone and the right key in 50% of the cases to generate a high tone (presented with 80% probabilities), while on few occasions, the left key generated a high tone and the right key generated a low tone instead (20% of the cases). For the Both condition (Figure 1C), participants were instructed to press the left key frequently, in 80% of the cases, and the right key rarely, in 20% of the cases, whereas the key–tone associations were the same as in the Intention condition. For all conditions, participants were made aware that, sometimes, another tone than the expected one will be presented and told to ignore it if that was the case and proceed normally to the next button press, as we wanted to avoid the possibility that participants stop throughout the block and report that something unexpected happened. Note that the key–tone associations as well as the frequently pressed key in Both condition have been counterbalanced and the above-described mappings reflect only one possibility. In all three conditions, participants' task was to press a key every approximately 1200 msec while, first, following the condition-specific instructions (press 50–50% or 80–20%) and, second, avoid producing fixed left/right

patterns of key presses (i.e., press the keys in a “random” sequence).

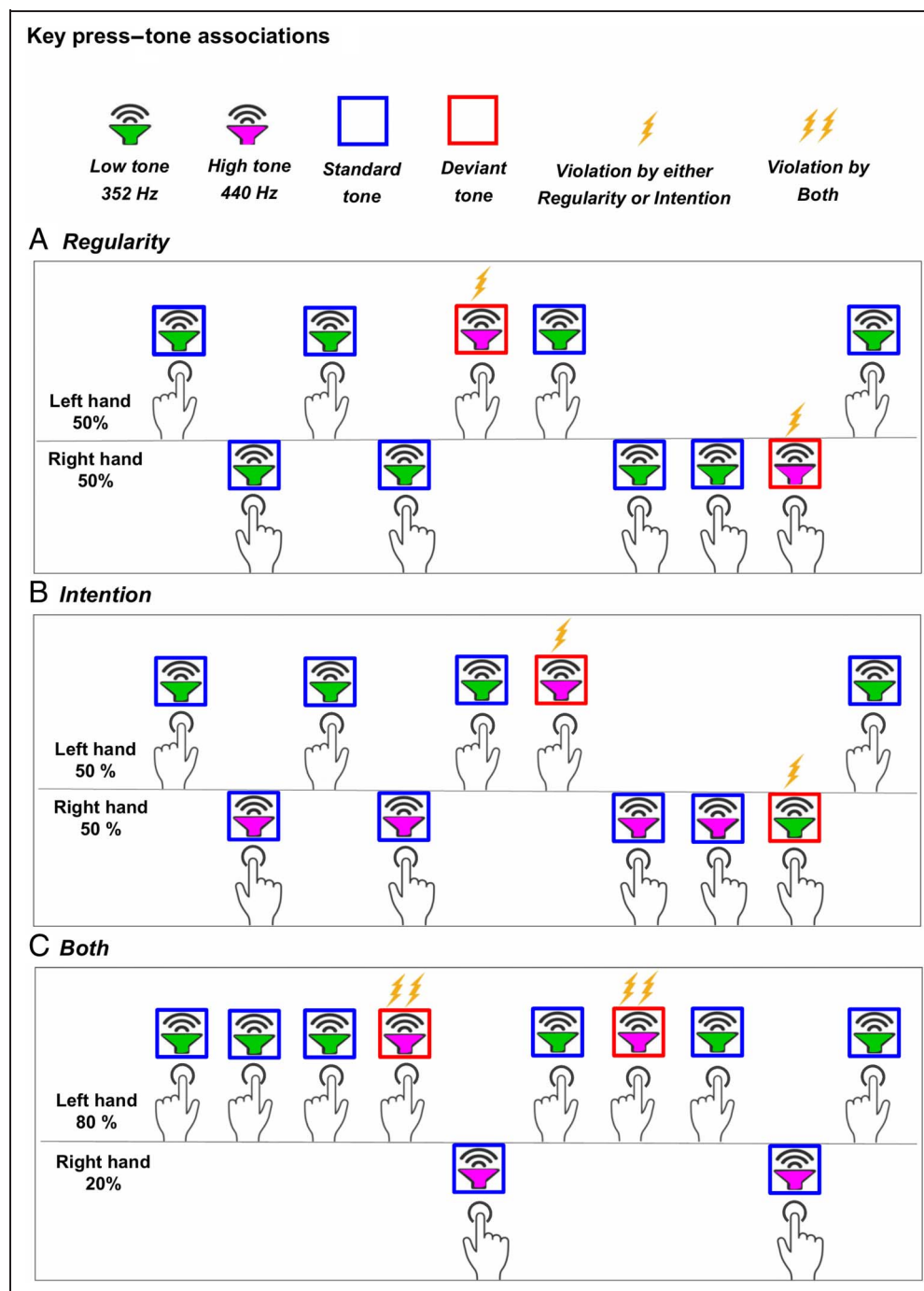
### Experimental Procedure

One session consisted of 10 experimental blocks. The duration of one block was about 3 min, and participants could take self-paced breaks in between. Three shorter practice blocks were completed at the beginning of every condition (blocks corresponding to the same condition were run one after another). The condition order as well as the key–tone associations and frequent versus rare key presses were counterbalanced between participants.<sup>1</sup> Two constraints were followed: first, Regularity and the frequently pressed key in Both generated the same standard tone, and second, the same key–tone association was kept between Intention and Both. We thus insured that there were no conflicting associations between conditions. One complete experimental session lasted for about 45 min.

The tone onset immediately followed the key press (with a delay of  $\sim 5$  msec, due to technical limitations). Trial duration was about 1200 msec, and participants were instructed to fixate on a fixation cross for the whole block duration. Figure 2 illustrates a possible sequence of trials at the start of a block, including screen feedback. The screen feedback was designed to help participants press the keys with equal (in Regularity and Intention) or unequal (Both) chances and to press a key about every 1200 msec. Each trial began with an indication of how many times the left versus right keys need to be pressed, displayed at the left and right sides of the fixation cross. This was indicated in numbers as well as in percentages (i.e., participants could see, for example, that the left key needs to be pressed 80 times, which represents 50% of the total number of key presses left for that block). The numbers on the left referred to the left hand and vice versa. Underneath the fixation cross and starting from the second trial, the timing between two consecutive key presses was displayed in milliseconds. Timing errors were defined as intervals shorter or longer by more than 400 msec than the indicated time (1200 msec). If a timing error occurred, a corresponding error message (“Too short/Too long”) was displayed on the screen, instead of the timing between the key presses—the tone was not presented in trials containing such errors, following which participants proceeded normally to the next button press. Note that because a tone was not presented, timing errors did not affect the total number of collected trials for the standard or deviant tones. Although the fixation cross was presented for the whole duration of one block, the screen feedback was updating every trial 600 msec after tone onset.

To have an indication of task compliance, we recorded timing errors (i.e.,  $\pm 400$  relative to 1200 msec) and fixed left–right sequences. The left–right sequences were analyzed online using the Walsh–Hadamard randomness test

**Figure 1.** An example of condition-specific key press–tone associations. In the Regularity condition (top), participants pressed the two buttons with 50–50% chances to generate a standard low tone (marked in green) with 80% probability (marked in blue square) and a deviant high tone (marked in pink) with 20% probability (marked in red square). In the Intention condition (middle), participants pressed the left button in 50% of the cases to generate a standard low tone and a deviant high tone and conversely the right button in 50% of the cases to generate a standard high tone and a deviant low tone. In the Both condition, the key press–tone associations were the same as in Intention, whereas the left key was pressed frequently (in 80% of the cases) and the right key rarely (in 20% of the cases). One lightning symbol marks tones that violate either regularity (top) or intention (middle). Two lightning symbols mark tones that violate both regularity and intention (bottom).

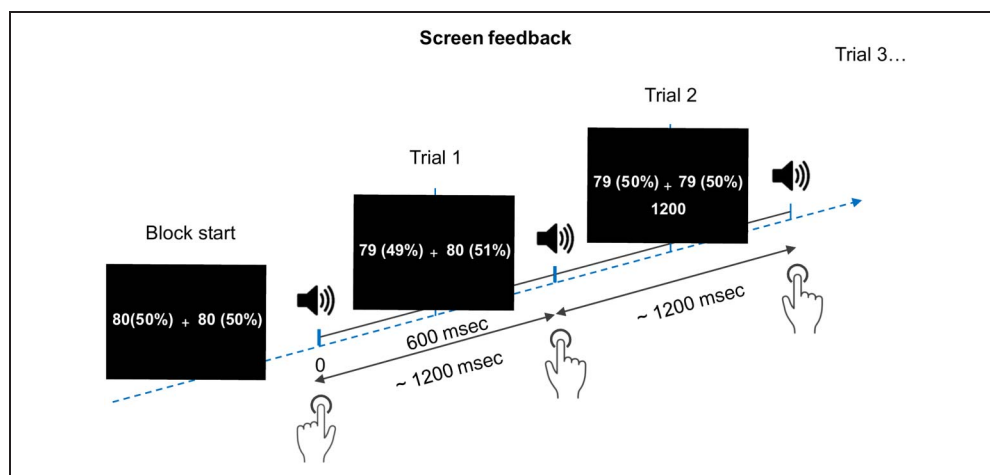


(Oprina, Popescu, Simion, & Simion, 2009), programmed in Octave within the experiment. The test uses a sequence of binary input (here, codes for left vs. right key presses) and detects randomness failure based on autocorrelation. If excessive fixed patterns were detected, participants were verbally warned at the end of the block. Note that measuring “pure” randomness was beyond our scope, because it is controversial whether humans can produce completely random sequences of actions, one major difficulty being the very definition of mathematical randomness (Wagenaar, 1972). This was rather implemented to make sure partici-

pants press the two keys equally often (or one key four times more often than the other one), without repeating a certain sequence excessively. The timing errors were analyzed offline as percentage errors from the total number of trials.

The Regularity condition contained 384 standard tones (80%) and 96 deviant tones (20%)—these were grouped in three blocks, each containing 160 trials, of which 80 corresponded to the left and 80 to the right key. In Intention, 192 high and 192 low tones inversely associated with the left and right keys were generated as

**Figure 2.** A possible sequence of trials. Every key press generates a tone (presented with a delay of  $\sim 5$  msec), followed by the screen feedback updating every 600 msec after tone onset. Participants fix their gaze on the fixation cross and press a key of their choice every about 1200 msec. The screen feedback indicates how many times the left versus right buttons need to be pressed in numbers and in percentages, and beginning with the second trial, the time interval between two consecutive key presses. Regularity and Intention blocks start with 80 trials for each of the left and right keys (50–50 chances). Both blocks start with 120 trials for the frequently pressed key and 30 trials for the rarely pressed key (80–20 chances).



standards (80% probability), and 48 low and 48 high tones inversely associated with the left and right keys were generated as deviants (20% probability). As in Regularity, these were presented in three blocks, each of 160 trials, of which 80 corresponded to the left and 80 to the right key. In Both, the frequently pressed key (80% of the cases) generated 384 standard tones (80% probability) and the 96 deviant tones (20% probability). The rarely pressed key (20% of the cases) generated 96 standard tones (80% probability) and 24 deviant tones (20% probability). These were presented in four consecutive blocks, each containing 150 trials, of which 120 trials corresponded to the frequent and 30 trials corresponded to the rare key. The standard–deviant sequence of tones within a block (and within the same key for Intention and Both) were randomized, with the constraint that the first two tones were always standards. Note that for the Both condition, we only analyzed the trials corresponding to the frequently pressed key. Thus, we recorded an equal number of trials for all three conditions: 384 standards versus 96 deviants.

### EEG Data Recording

EEG data were continuously recorded at a sampling rate of 500 Hz with a system equipped with 64 Ag–AgCl active electrodes, using a BrainAmp amplifier and the Vision Recorder software (Brain Products GmbH, Munich, Germany). Fifty-eight electrodes were mounted in an elastic cap (actiCAP) following the extended international 10–20 system (Chatrian, Lettich, & Nelson, 1985). Two additional electrodes were placed on the mastoids. One electrode placed on the tip of the nose served as online reference, a ground electrode was placed on the forehead, and three elec-

trodes were used to record EOG activity, two of which were placed on the left and right outer canthi and one below the left eye.

### EEG Preprocessing

The preprocessing was carried out in three steps using the EEGLAB MATLAB-based software (Delorme & Makeig, 2004). First, data were filtered using a 0.1-Hz high-pass and 45-Hz low-pass windowed sinc finite impulse response filter (Hamming window, filter order 8250 [high pass] and 166 [low pass]), in accordance with the recommendations of Widmann, Schröger, and Maess (2015). On average, 1.29 channels containing extreme amplitudes were removed using a deviation criterion that “calculates the robust  $z$  score of the robust standard deviation for each channel” (Bigdely-Shamlo, Mullen, Kothe, Su, & Robbins, 2015). Data were then epoched around the tone presentation ( $-200$  to  $600$  msec). Epochs with amplitudes exceeding a  $600\text{-}\mu\text{V}$  Delta threshold were removed. Second, an independent component analysis was computed on the raw data, which were first filtered with a 1-Hz high-pass and 45-Hz low-pass filter, epoched ( $-200$  to  $600$  msec relative to tone presentation) and cleaned by removing the same bad channels and epochs detected at the first step. The obtained weights were stored and transferred to the data sets obtained at the first step. Third, the removal of components containing eye-related (blinks, lateral eye movements) and muscle artifacts was done by visual inspection and based on measures computed with FASTER (Nolan, Whelan, & Reilly, 2010), ADJUST (Mognon, Jovicich, & Bruzzone, 2011), and SASICA (Chaumon, Bishop, & Busch, 2015).<sup>2</sup> The missing channels were interpolated using the built-in EEGLAB spherical interpolation function, and data were baseline

corrected using the 200-msec prestimulus interval. Epochs with amplitudes still exceeding a 200- $\mu$ V Delta threshold after the independent component analysis cleaning were removed—epochs removed at both the first and third steps represented less than 1% from the total number of trials. Finally, condition-specific averages were calculated.

### PCA Analysis

A temporal PCA was performed using the ERP PCA toolkit MATLAB-based toolbox (Dien, 2010). We computed the PCA on the individual averages of all conditions using Promax rotation ( $k = 3$ ) with a covariance relationship matrix and Kaiser weighting. Horn's parallel test was used to determine the number of components to be retained.

### Statistical Analysis

Each component of interest identified by temporal PCA was separately tested using first frequentist and second Bayesian repeated-measures ANOVA (rANOVA); the analyses were conducted using IBM SPSS Statistics 25 and JASP 0.9.1.0, respectively. Note that for the Intention condition, standards from the left and right hands were pooled together for the analyses (irrespective of whether they were low- or high-pitched)—similarly, the deviants from both hands were analyzed together. For the Both condition, only tones generated by the frequently pressed key were analyzed. A  $3 \times 2$  design with factors Condition (Regularity vs. Intention vs. Both)  $\times$  Stimulus (standard vs. deviant) was used for the frequentist analysis. Statistical significance was defined at the .05 alpha level, and results are reported including the eta-square effect sizes ( $\eta^2$ ). Follow-up  $t$  tests were computed for statistically significant interactions. The complementary  $3 \times 2$  Bayesian analysis was calculated to test all alternative models, including main effects and interactions against the null model, which included only the random factor, that is, participants' variation. The Bayes factor ( $BF_{10}$ ) was calculated using 10,000 sample repetitions; the null hypothesis corresponded to a standardized effect size  $\delta = 0$ , and the alternative hypothesis was defined as a Cauchy prior distribution centered around 0 (Rouder, Morey, Speckman, & Province, 2012). Bayesian  $t$  tests followed up on the effects of the models including interactions, provided these supported the alternative hypothesis best. Lastly,  $BF_{Inclusion}$  calculated across matched models (i.e., models that include vs. do not include the effect) provided a measure of change odds from prior to posterior distributions. These were only calculated if more than one model supported the alternative hypothesis to have clear evidence whether the main effects or the interaction explain the data best. In accord with existing recommendations on how to interpret the Bayes factor (Lee & Wagenmakers, 2014; Jeffreys, 1961), values  $\leq 0.3$  were taken as evidence in favor of the null hypothesis, values  $\geq 3$  as

evidence in favor of the alternative hypothesis, whereas values close to 1 were considered poor evidence.

Finally, we checked by means of  $t$  tests the Regularity + Intention versus Both additivity model (e.g., see Pieszek et al., 2013, for a similar procedure). For this purpose, the component scores representing the differences between deviants and standards (deviant – standard) were calculated for every condition, the difference scores for Regularity and Intention being subsequently added together. Note that the additivity model was only tested if sensory prediction error effects (significant differences between deviants and standards) were found in all three conditions and only for those sensory components in which such effects were found.

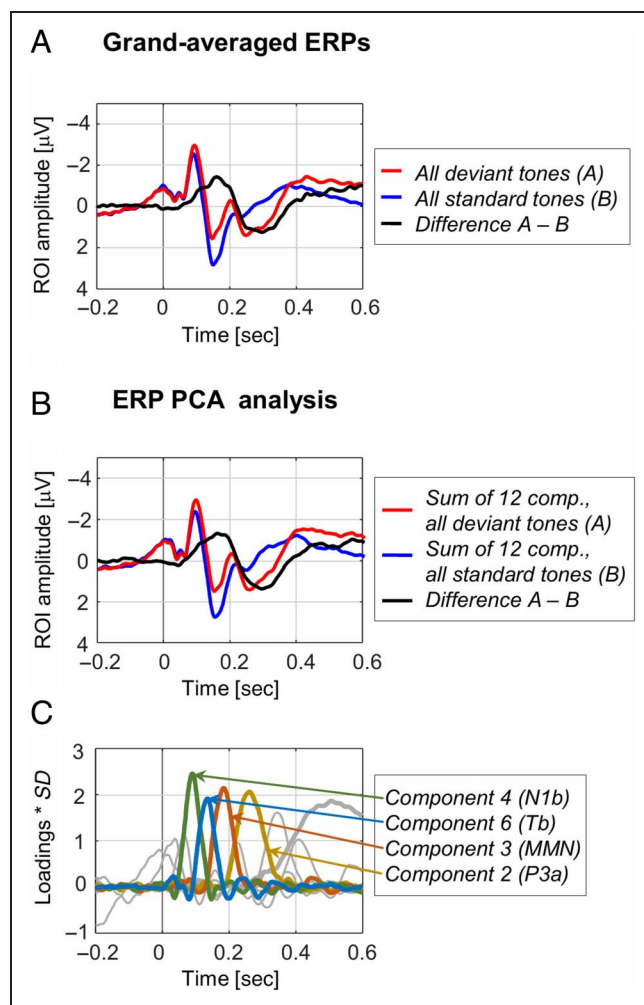
## RESULTS

### Timing Errors

The key press-to-key press time intervals longer or shorter than 1200 msec by more than 400 msec were recorded as timing errors. We calculated these as error percentages (% ERR) relative to the total number of trials contained in every block—blocks corresponding to the same condition were subsequently averaged. Participants made on average 2.76 %ERR in Regularity ( $SD = 3.51$ , range = 0–12.29), 3.29 %ERR in Intention ( $SD = 3.40$ , range = 0–11.88), and 2.38 %ERR in Both ( $SD = 3.28$ , range = 0–10)—these indicate they followed the instructions and pressed the keys at the suggested pace. A one-way ANOVA including the three conditions (Regularity, Intention, Both) revealed no significant Condition differences,  $F(2, 46) = 2.78$ ,  $p = .114$ ,  $\eta^2 = .090$ . Committed errors do not indicate fewer “correct” trials for the analysis of ERPs, because for the trials containing timing errors, no tone was generated.

### ERP PCA Results

Figure 3 displays the grand-averaged ERPs, along with the PCA results. According to Horn's parallel test, 12 components were extracted explaining over 95% of the total epoch variability (Figure 3C). The stimulus-specific waves represented by the sum of the 12 retained components firmly correspond to the stimulus-specific grand-averaged ERPs—see Figure 3A–B for a visual comparison at the level of an ROI representing an average of Fz, FCz, and Cz, electrodes that typically are of interest in auditory processing (e.g., see results in Timm et al., 2014; Hughes et al., 2013; SanMiguel, Saupe, et al., 2013; Horváth et al., 2008; Näätänen et al., 2005). Of the retained 12, we focused our attention on four, Components 2, 3, 4, and 6, presumably representing the P3a, MMN, N1b, and Tb peaks, respectively. The selection of the four components of interest was based on latency and topographical information. Note that they are ordered not by chronological peak latency but by the explained variance. That is, Component 2 corresponding to the P3a peak at 260 msec explains  $\sim 15.2\%$  of the epoch variability. Component 3



**Figure 3.** ERP PCA results. (A) Grand-averaged ERPs are presented for the deviant (red) and standard (blue) tones, along with the difference wave (black), averaged across all three conditions, for a ROI composed of the electrodes Fz, FCz, and Cz. (B) Following the PCA analysis, 12 components explaining over 95% of the epoch variability were retained—the waves representing the sum of these 12 components for the deviant (red) and standard (blue) tones, as well as the difference wave (black), firmly correspond to the original grand-averaged ERPs. (C) The 12 retained components are presented individually. Of these, Components 2, 3, 4, and 6 corresponding to the P3, MMN, N1b, and Tb peaks, respectively, were further analyzed—these are marked in color. Component 1, marked in the thicker gray line, is related to the reorienting negativity or N3.

corresponding to the MMN response reflecting an increase in negativity for the deviant compared with standard tones peaking at 184 msec (P2 range) explains  $\sim$ 9.7% of the epoch variability. Component 4 representing the sensory-specific N1b peak at 90 msec and Component 6 representing the Tb peak (corresponding to the T-complex) at 134 msec explain  $\sim$ 4.6% and  $\sim$ 2.6% of the epoch variability, respectively. Note that the identified time courses and topographies (i.e., early latency and frontocentral distribution for N1b, later latency, and temporal distribution for Tb) correspond to previous studies reporting N1-constituent components (SanMiguel, Todd, et al., 2013;

Näätänen & Picton, 1987). Component 1 (not analyzed here but marked in the thicker gray line in Figure 3C) explains about half of the whole epoch variability (51.3%) and peaks at 506 msec—this presumably represents the reorienting negativity or N3 peak (Kotchoubey & Pavlov, 2019). The time-invariant component scores represent the contribution of each component of interest to the ERP wave—these have been subjected to statistical analyses. The time-variant loadings of the components reflect their contribution to the voltage maps at each point in time.

For each component of interest, we analyzed component scores at the electrodes showing the largest score activations. Correspondingly, the N1 scores were analyzed at electrode Fz. The Tb has a bilateral distribution peaking around the T7 and T8 temporal electrodes—we thus analyzed the average of the component scores corresponding to the two electrodes. Finally, the scores corresponding to MMN and P3a components were analyzed at electrode Cz. Figures 4 and 5 display for each component the condition-specific activations for the standard and deviant tones, along with the corresponding topographical maps (N1b, Figure 4A; Tb, Figure 4B; MMN, Figure 5A; P3a, Figure 5B). Violin plots in Figure 6 display the condition-specific effects (i.e., component scores plotted as deviant–standard differences) for each of the four components (N1b, Figure 6A; Tb, Figure 6B; MMN, Figure 6C; P3a, Figure 6D). We further report component-specific statistical results. Main effects and interactions obtained in the frequentist versus Bayesian analyses are summarized in Table 1.

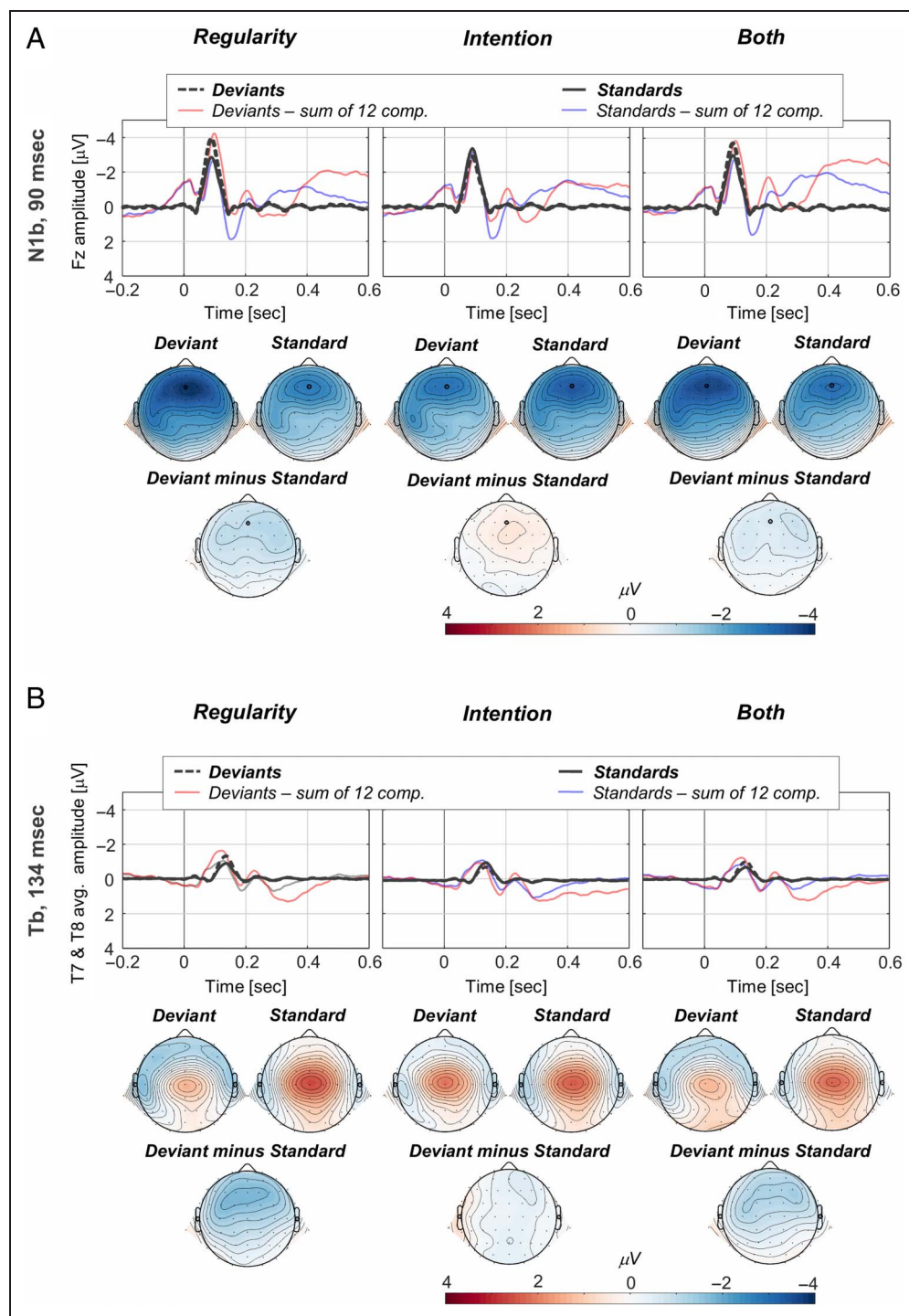
#### *N1b Enhancement for Deviants in Regularity and Both, but Not in Intention*

The frequentist rANOVA revealed a main effect of Stimulus,  $F(1, 23) = 5.92, p = .023, \eta^2 = .205$ , and an interaction of Condition  $\times$  Stimulus,  $F(2, 46) = 18.86, p < .001, \eta^2 = .451$ . Follow-up  $t$  tests indicate the N1b component scores are significantly enhanced for the deviant tones in Regularity,  $t(23) = 4.39, p < .001$ , and Both,  $t(23) = 2.76, p = .011$ , but not in Intention, where, in contrast, the scores for the standard tones are enhanced,  $t(23) = -2.26, p = .034$ . The Bayesian rANOVA favored the full model, including the main effects and the interaction term (Condition + Stimulus + Condition  $\times$  Stimulus,  $BF_{10} = 3.61 \pm 1.96\%$ ; see Table 1), whereas the models containing main effects of Condition and Stimulus only provided anecdotal evidence. Follow-up Bayesian  $t$  tests mirrored the frequentist results by providing strong evidence for the alternative hypothesis in Regularity ( $BF_{10} = 138.77 \pm <0.001\%$ ), moderate in Both ( $BF_{10} = 4.47 \pm <0.001\%$ ), and only anecdotal evidence in Intention ( $BF_{10} = 1.78 \pm 0.005\%$ ).

#### *Tb Enhancement for Deviants in Regularity, but Not in Intention and Both*

The frequentist rANOVA revealed a significant main effect of Stimulus,  $F(1, 23) = 5.44, p = .029, \eta^2 = .191$ , and an

**Figure 4.** N1b and Tb PCA results. (A) The N1b component peaks at 90 msec and is largest at the Fz electrode. (B) The Tb component peaks at 134 msec and is largest over temporal T7 and T8 electrodes, the displayed waves representing a mean of the two. For all three conditions, the component-specific waves for the standards (dark full lines) and deviants (dark dashed lines) are displayed along with the “reconstruction waves,” representing the sum of the 12 retained components for the deviants (faded red lines) and standards (faded blue lines). The topographical maps have been calculated based on spherical spline interpolation and illustrate the deviants and standards evoked responses, as well as the deviants–standards difference maps. The electrodes marked on the topographical maps (N1b→Fz, Tb→T7, and T8) represent the ones included in the analysis.



interaction of Condition  $\times$  Stimulus,  $F(2, 46) = 3.33$ ,  $p = .044$ ,  $\eta^2 = .127$ . Follow-up  $t$  tests indicate the Tb component scores are significantly enhanced for the deviant tones in Regularity,  $t(23) = 2.92$ ,  $p = .008$ , whereas in Both, a nonsignificant trend was observed,  $t(23) = 1.77$ ,  $p = .089$ . In Intention, the difference between the standard and deviant tones was not significant,  $t(23) = 0.25$ ,  $p = .798$ . The Bayesian rANOVA favored the model containing the main effects (Condition + Stimulus,  $BF_{10} = 3.22 \pm 1.36\%$ ; see Table 1), whereas all other models,

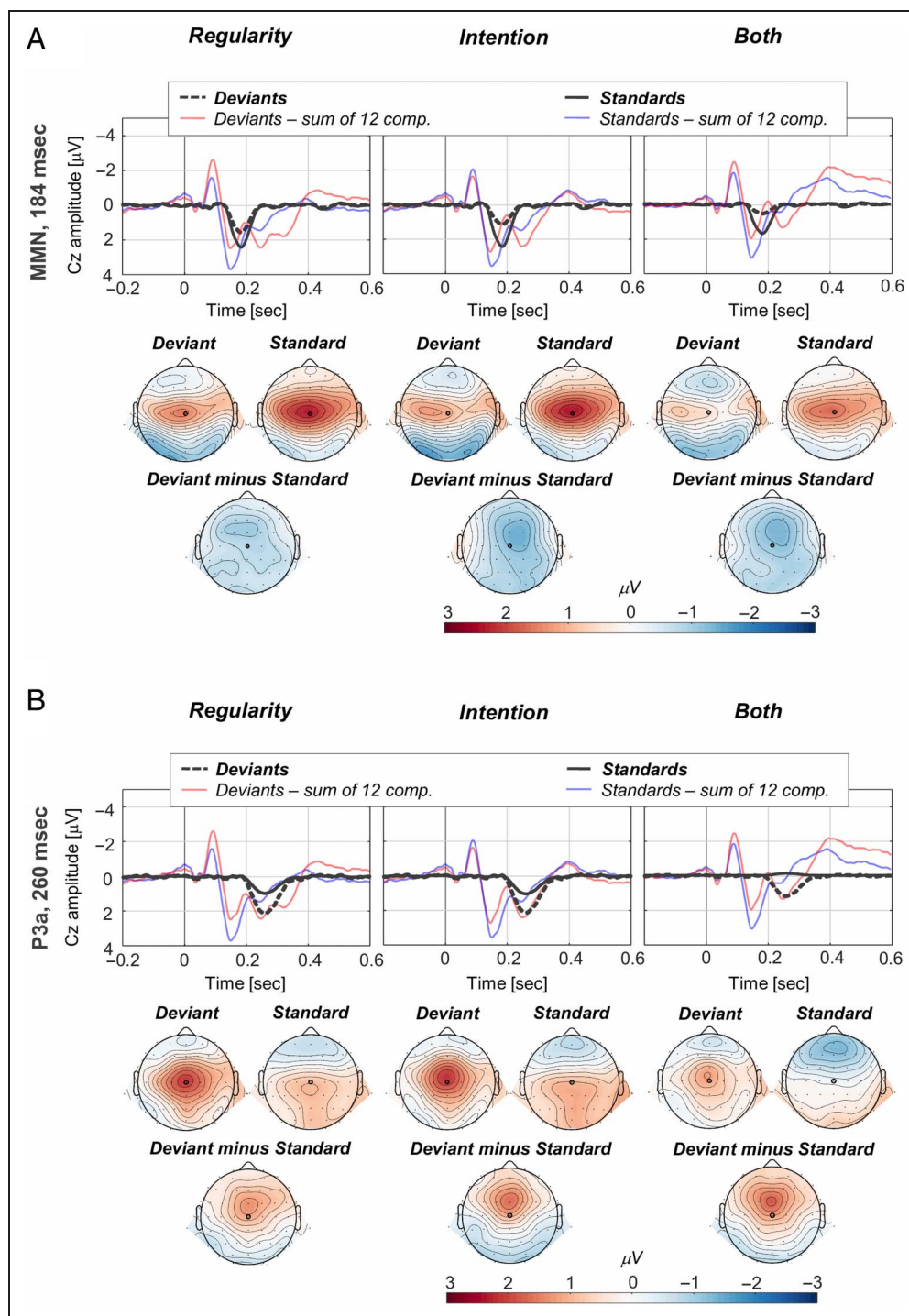
including the one with the interaction term, only provided anecdotal evidence. The Bayesian analysis therefore did not bring conclusive evidence in favor or against the alternative hypothesis containing the interaction term ( $BF_{10} = 0.96 \pm 3.45\%$ ).

#### MMN in Regularity, Intention, and Both

The frequentist rANOVA revealed a main effect of Stimulus,  $F(1, 23) = 11.24$ ,  $p = .003$ ,  $\eta^2 = .328$ , with smaller



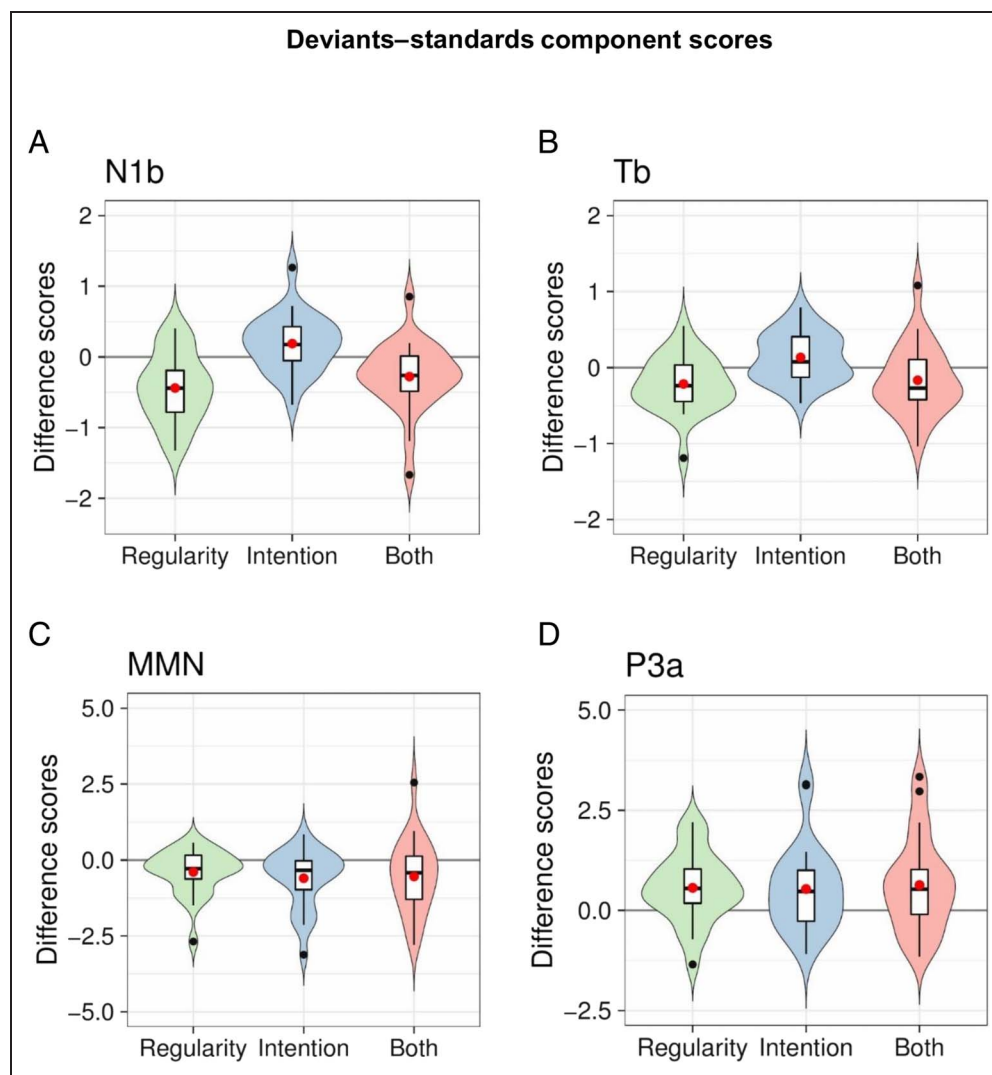
**Figure 5.** MMN and P3a PCA results. The MMN and P3a components are largest at electrode Cz. MMN peaks at 184 msec (A), and the P3a peaks at 260 msec (B). For all three conditions, the component-specific waves for the standards (dark full lines) and deviants (dark dashed lines) are displayed along with the “reconstruction waves” representing the sum of the 12 retained components for the deviants (faded red lines) and standards (faded blue lines). The topographical maps have been calculated based on spherical spline interpolation and illustrate the deviants and standards evoked responses, as well as the deviants–standards difference maps. The electrode marked on the topographical maps (Cz) represents the one included in the analysis.



positive scores for the deviant ( $M = 0.50$ ) as compared with the standard ( $M = 1.01$ ) tones. A trend approaching significance was observed as a main effect of Condition,  $F(2, 46) = 3.10, p = .054, \eta^2 = .119$ , but no significant interaction term was observed,  $F(2, 46) = 0.43, p = .651, \eta^2 = .018$ ; the presence of each of the three MMNs was confirmed by paired  $t$  tests: Intention,  $t(23) = 3.06, p = .006$ ; Regularity,  $t(23) = 2.63, p = .015$ ; Both,  $t(23) = 2.19, p = .038$ ). Thus, the MMN is present across all three conditions, but not significantly different between condi-

tions. The Bayesian rANOVA favored the model containing the main effects of condition and stimulus (Condition + Stimulus,  $BF_{10} = 316.99 \pm 1.24\%$ ; see Table 1), whereas the models containing the main effect of stimulus and the full model also brought evidence in favor of the alternative hypothesis. However, averaged across the matched models, the  $BF_{\text{Inclusion}}$  suggests only the main effect of Stimulus should be retained ( $BF_{\text{Inclusion}} = 246.06$ ), whereas the main effects of Condition and the interaction only provided anecdotal evidence for the alternative hypothesis or

**Figure 6.** Condition-specific effects: deviants–standards. Violin plots display the condition-specific deviant–standard component scores for N1b (A), Tb (B), MMN (C), and P3a (D) components. The estimated density distributions (displayed in green for Regularity, blue for Intention, and red for Both) are shown along with boxplots indicating the medians, interquartile ranges, and confidence intervals, whereas the means are displayed in red dots. The black dots represent individual data points falling outside the confidence intervals. Note that the scores are component-specific and the scales between A and D do not correspond to each other.



moderate evidence for the null hypothesis, respectively (Condition:  $BF_{\text{Inclusion}} = 1.59$ ; Condition  $\times$  Stimulus:  $BF_{\text{Inclusion}} = 0.14$ ). Therefore, mirroring the frequentist results, this indicates that there are no condition differences for the MMN component.

#### *P3a Enhancement for Deviants in Regularity, Intention, and Both*

The frequentist rANOVA revealed a significant main effect of Stimulus,  $F(1, 23) = 12.14, p = .002, \eta^2 = .346$ , with larger positive scores for the deviant ( $M = 0.87$ ) as compared with standard ( $M = 0.30$ ) tones, and a significant main effect of Condition,  $F(2, 46) = 6.64, p = .003, \eta^2 = .224$ , with larger positive values for Intention ( $M = 0.76$ ), followed by Regularity ( $M = 0.74$ ) and Both ( $M = 0.26$ ). No significant interaction term was observed,  $F(2, 46) = 0.09, p = .908, \eta^2 = .004$ , suggesting the P3a enhancement effect is present across all three conditions, but not significantly different between conditions. The Bayesian rANOVA favored the model containing the main effects of Condition

and Stimulus (Condition + Stimulus,  $BF_{10} = 18231.89 \pm 0.95\%$ ; see Table 1), whereas all other models containing the main effect of Condition and the main effect of Stimulus, as well as the full model, also brought evidence in favor of the alternative hypothesis. However, averaged across the matched models, the  $BF_{\text{Inclusion}}$  suggests the models containing the main effects of Condition ( $BF_{\text{Inclusion}} = 23.17$ ) and stimulus ( $BF_{\text{Inclusion}} = 1719.02$ ) should be retained, whereas the interaction provided moderate evidence for the null hypothesis ( $BF_{\text{Inclusion}} = 0.12$ ). Again, mirroring the frequentist results, this suggests that there are no condition differences for the P3a component.

#### *No Additivity of Regularity and Intention Effects*

We calculated the Regularity + Intention versus Both additivity model for the MMN component, where significant prediction errors were observed in all three conditions. The Regularity + Intention difference scores do not equal the Both difference scores (i.e., are significantly different,  $t(23) = -2.08, p = .048$ ), with the former being

**Table 1.** Results of Statistical Analyses

Comp.	Frequentist Effects				Bayesian Models		
		<i>F</i>	<i>p</i>	$\eta^2$		<i>BF</i> <sub>10</sub>	± %ERR
N1b	Cond	0.40	.669	.017	Cond	0.11	0.70
	<b>Stim</b>	<b>5.92</b>	<b>.023</b>	<b>.205</b>	Stim	1.79	0.85
					Cond + Stim	0.21	2.09
	<b>Cond × Stim</b>	<b>18.86</b>	<b>&lt;.001</b>	<b>.451</b>	<b>Cond + Stim + Cond × Stim</b>	<b>3.61</b>	<b>1.96</b>
Tb	Cond	2.38	.104	.094	Cond	1.54	0.87
	<b>Stim</b>	<b>5.44</b>	<b>.029</b>	<b>.191</b>	Stim	1.81	1.21
					<b>Cond + Stim</b>	<b>3.22</b>	<b>1.36</b>
	<b>Cond × Stim</b>	<b>3.33</b>	<b>.044</b>	<b>.127</b>	Cond + Stim + Cond × Stim	0.96	3.45
MMN	Cond	3.10	.054	.119	Cond	1.09	0.65
	<b>Stim</b>	<b>11.24</b>	<b>.003</b>	<b>.328</b>	Stim	199.07	1.05
					<b>Cond + Stim</b>	<b>316.99</b>	<b>1.24</b>
	Cond × Stim	0.43	.651	.018	Cond + Stim + Cond × Stim	46.24	2.74
P3a	<b>Cond</b>	<b>6.64</b>	<b>.003</b>	<b>.224</b>	Cond	10.06	0.50
	<b>Stim</b>	<b>12.14</b>	<b>.002</b>	<b>.346</b>	Stim	786.12	0.79
					<b>Cond + Stim</b>	<b>18231.89</b>	<b>0.95</b>
	Cond × Stim	0.09	.908	.004	Cond + Stim + Cond × Stim	2285.46	1.69

For all analyzed components, the frequentist main effects and interactions with their corresponding *p* values and effect sizes are displayed (left), along with the Bayes factors and corresponding errors for the full models including main effects and interactions (right). The significant main effects and interactions from the frequentist analyses as well as the models with the best explanatory power from the Bayesian analysis are highlighted in **bold**.

more negative (i.e., larger effects,  $M = -0.99$ ) than the latter ( $M = -0.54$ ). Thus, an additivity model does not suit these data.

## DISCUSSION

Action predictions based on action intention and sensory predictions based on tone regularity are often described as similar mechanisms, but a direct comparison is yet missing. We addressed this issue by using a “self-generation oddball paradigm,” where participants produced standard and deviant tones (high or low pitched) by performing “random” sequences of left and right button presses. By manipulating the action–tone association as well as the likelihood of performing one action over the other, we contrast predictions based on tone regularity versus intention to produce a specific tone versus both intention and regularity. Our results indicate that the N1b and Tb components of the N1 response are modulated by regularity violations, but not by intention violations. Intention and regularity violations are reflected in the MMN response,

which importantly is elicited with and without high global probability of the standard tone. Even though regularity and intention might represent independent generative sources, their resulting prediction errors seem to integrate, rather than add up—that is, we did not observe stronger effects (indicating additivity) when the two were present together. Finally, similar P3a effects for all conditions suggest that, regardless of whether the sensory predictions are implemented based on either tone regularity, motor intention, or both, this does not influence the deviance detection mechanism further implemented at the next processing step. As follows, we discuss these findings in more detail.

### Bottom–Up Regularity-based Prediction Errors Are Reflected in the N1 and MMN Components

The use of temporal PCA has led to the distinct identification of the N1b and Tb N1-constituent components, as well as of the MMN component. We found N1b enhancement effects for the deviant relative to the standard tones in the Regularity and Both conditions, but not in the Intention condition—this was supported by frequentist

as well as Bayesian analyses. Additionally, we found Tb enhancement for the deviant relative to the standard tones in the Regularity condition only. For the MMN component, the typical enhancement for the standard relative to the deviant tones was observed (peaking at P2 latency range), leading to negativity responses for the deviant–standard evoked tones. This effect was present across the Intention, Regularity, as well as Both conditions, with no differences between the three, as indicated by frequentist as well as Bayesian analyses.

The N1b and Tb effects presumably reflect stimulus-specific adaptation of the neuronal responses (Grill-Spector, Henson, & Martin, 2006). The effect sizes and Bayes factors suggest the N1 magnitude decreases as a function of global tone probability, with strong effects in the Regularity condition, followed by a decrease in Both and finally no effects in Intention, where global regularity (standard-to-deviant probability, regardless of action) is absent. That is, in the Regularity condition, the standard tone was overall presented in 80% of the cases, because the two buttons were equally pressed and generated the same frequent (and infrequent) tone. In the Both condition, the two buttons were pressed with 80% versus 20% chances and were inversely associated with the two tones. Subsequently, here, the frequently presented tone was the standard generated by the frequently pressed key, with a global regularity of 68% (resulting from frequent key → standard 64% vs. deviant 16%; infrequent key → standard 16% vs. deviant 4%; standard 64% + deviant 4% [same tone between the two keys] = 68%). Finally, in the Intention condition, the mapping of standards and deviants was inversely associated with the left and right keys, which were pressed equally frequent, meaning that the two tones were overall presented with equal chances.

The distinction typically made between the N1-constituent components (SanMiguel, Todd, et al., 2013; Näätänen & Picton, 1987; McCallum & Curry, 1980; Wolpaw & Penry, 1975) indicates that these N1b and Tb results can be interpreted as a consequence of “true” sensory predictions, in contrast to mere orienting responses captured by the “unspecific” N1, elicited with long ISIs (SanMiguel, Todd, et al., 2013). Indeed, unlike most self-generation studies, we used a short tone-to-tone interval; we made sure the timing between two consecutive button presses was stable around 1200 msec across all trials (in comparison to a range in between 2000 and 6000 msec, in typical self-generation studies). Participants proved to be able to keep the correct pace for all blocks and conditions, as indicated by the few timing errors (less than 3.5% on average), when they produced intervals longer or shorter than 1200 msec by more than 400 msec.

Next, our data suggest that the N1 component is followed by MMN responses in the Regularity and Both conditions. The N1–MMN succession is a typical result associated with regularity-based prediction errors, where while N1 reflects stimulus adaptation, the MMN

represents a memory- or prediction-driven comparison of the expected versus received input (Garrido et al., 2009; Näätänen et al., 2005). This pattern of results is also compatible with the proposed stages of auditory distraction (Horváth et al., 2008), where, at an initial sensory processing step, the N1 represents first-order and the MMN second-order change detectors. Thus, in the Regularity and Both conditions, first- and second-order prediction errors are implemented in a bottom–up manner via global tone regularity.

### **Top–Down Intention-based Prediction Errors Are Reflected in the MMN Component**

As already mentioned, the MMN is present in the Regularity and Both conditions, where the global likelihood of the standard tones was higher than the one of the deviants, but also in the Intention condition, where the two tones were presented with equal chances. Because this effect is robust with and, importantly, without global tone regularity, we propose that it represents an intention-based prediction error (elicited in the Intention condition). Consequently, we show that the intention-based MMN is implemented in a top–down manner, when controlled for neural adaptation reflected in the early N1 response (Jacobsen & Schröger, 2001). One important distinction between adaptation at low levels due to regular input and top–down effects is that the first is an automatic and necessary side effect of bottom–up sensory processing, whereas the second involves high-level expectations, which are fed back down the cortical hierarchy to achieve effects at sensory levels (Lee & Mumford, 2003). It has indeed been shown that top–down expectations regarding the quality (high or low) of individual tones within a sequence modulate the sensory ERP components starting from 100 msec (Widmann et al., 2004). Altogether, these results indicate that intention and regularity, in line with earlier (but untested) suggestions, are distinct “sources for a single mechanism” (Lange, 2013).

Referring to the intention-based expectations on deviance processing, Waszak and Herwig asked participants to generate standard versus deviant tones by voluntary left and right key presses, similarly to here. Unlike here, in their design, both key presses generated the same standard and deviant tones in a test phase, whereas in a previous acquisition phase, the left versus right actions have been associated with either the standard or the deviant tone, with 100% certainty. Thus, in the test phase, based on the intention to press a key over the other one, either the standard or the deviant tone were to be expected, but based on tone regularity, the same standard tone was frequently presented, regardless of the chosen action. They have found P3a effects between the predicted and unpredicted deviants (i.e., the deviants associated with the same button press as in the acquisition phase vs. the deviants associated with the button press that

has in the acquisition phase been associated with the other tone) and conclude that this P3a effect is due to previously formed intention-based expectations (Waszak & Herwig, 2007). Our results go a step further and show that, when controlled for stimulus regularity, the intention-based expectations modulate deviance processing even earlier at sensory levels reflected in the MMN response.

Congruent with the present MMN intention result, a recent study by LeBars and colleagues reported that effects around 200 msec (at the level of what is described as the N2b component) do indeed depend on whether the participants are able to choose or not which button to press. Specifically, similar to here, they had participants generate low and high standards and deviants, which were inversely associated with left and right button presses, and showed that only when the choice of which key to press was determined by participants' intention, in contrast to externally cued, mismatch answers were elicited (Le Bars, Darriba, & Waszak, 2019).

### Further Deviance Detection Is Reflected in the P3a Response for Bottom-Up and Top-Down Predictions

Following up on the stages of auditory distraction, if the deviation detection reflected in the N1 and MMN components at the sensory processing step exceeds a certain threshold, a second processing step reflected in the P3a response follows (Horváth et al., 2008). We found P3a effects (larger positive amplitudes for the deviant, compared with the standard tones) in all three conditions, with no differences between conditions—these effects were once again supported by frequentist as well as Bayesian analyses. First, this indicates that the deviance has been strongly perceived in all three conditions pointing to the success of the experimental manipulation. Second, the P3a results point out that, regardless of whether the predictions at the initial sensory processing levels have been violated based on either Regularity, Intention, or Both, this does not seem to influence the following processing level where the change detection mechanism reflects an involuntary attentional switch toward motivational (i.e., expectancy violating) stimuli (Nieuwenhuis et al., 2011).

In the light of Waszak and Hervig's P3a interpretation (Waszak & Herwig, 2007), we cannot rule out the possibility that, in fact, the P3a effects reported here might also reflect intention-based signatures across conditions. That is, given that participants always chose when and which key to press to produce a tone (regardless of whether that tone was hand specific in the Intention and Both conditions or associated with both hands in the Regularity condition), the intention per se to perform an action to produce an effect must have been a factor in all three conditions similarly. Therefore, to better understand the P3a effect in this context, it would be necessary to reduce the intention-related processes in the case of regularity-based predictions, for example, by cuing the left and right button

presses, similar to Le Bars et al. (2019). However, Le Bars et al. (2019) do not show to what extent cuing the actions, in contrast to intentional action, affects the magnitude of the P3a effect. It thus remains for future research to establish the precise functional interpretation of the P3a component following sensory predictions.

### Concurrent Violations of Intention and Regularity Do Not Lead to Stronger Prediction Errors

The interaction of top-down and bottom-up information is widely discussed in the literature. It is generally assumed that, in the context of hierarchical processing of predictive information, bottom-up information and top-down expectations are constantly contrasted and integrated in cortical feedforward/feedback loops (Lee & Mumford, 2003). We did not find stronger prediction errors elicited by concurrent violations of regularity and intention—that is, the effects we report for the sensory N1 and MMN components are not larger in the Both condition, which would suggest additivity of Regularity and Intention. This conclusion is supported by frequentist and Bayesian analyses, as well as by testing the additivity model (i.e., Regularity + Intention vs. Both). These data therefore suggest that convergent predictions by regularity and intention are integrated at lower levels of processing (as opposed to being represented independently and eliciting separate prediction error responses), despite the two presumably having independent generative models.

A relevant study by Pieszek and colleagues compared bottom-up regularity predictions with top-down predictions based on audio-visual pairing. Specifically, within a trial, the two types of information could either be contradictory (i.e., one confirming vs. the other violating the tone expectation), or both confirming, or concurrently violating the expectation. Similar to this study, they showed that bottom-up and top-down violations individually lead to prediction errors, as expressed by MMN and incongruency responses, respectively. However, unlike here, they report an additive bottom-up + top-down model, where the difference wave corresponding to the concurrent violations roughly matches the sum of the two difference waves corresponding to the independent predictions (Pieszek et al., 2013).

On the one hand, according to the predictive coding theory, the prediction mechanism generates predictions regarding both the context of the incoming stimulation, as well as about the expected precision (Feldman & Friston, 2010)—these effects can, in turn, be mediated by attention (Schröger et al., 2015). From this perspective, it would make sense that more precision (i.e., bottom-up + top-down) would lead to stronger prediction errors, provided the expectations are violated—like in Pieszek et al. (2013), for example. On the other hand, it is also likely that the mechanism works based on an “efficiency rule,” where, in the absence of special attentional resources, once a reliable source of information is available for the

prediction (extracted from either tone regularity or action intention), additional sources become redundant.

Regarding the first hypothesis according to which more precision should enhance the prediction error, it is possible that extra attentional resources would be required to enhance the prediction errors by additional information (as opposed to additional information being redundant). It is generally believed that attention and prediction work together to enhance precision (Schröger et al., 2015), and this effect could also be transferable to concurrent bottom-up and top-down predictions. To conclude, it remains for future research to establish whether including an attention manipulation in studying the nonconflicting predictions based on regularity and intention changes the prediction precision and leads to higher prediction errors for concurrent (both regularity and intention) in comparison to single violations (either regularity or intention).

The second hypothesis becomes likely if we consider that, in our design, throughout blocks corresponding to the same condition (and within individual trials), the bottom-up versus top-down predictions did not contradict each other, in contrast to the study by Pieszek et al. (2013). Specifically, in their design, within the same trial, individual predictions could be violated based on tone regularity, but confirmed based on the visual-auditory pairing or vice versa. This situation is confusing, in comparison to when both information types unanimously confirm or violate expectations; thus, it is unsurprising that concurrently violating both predictions (i.e., high certainty) leads to a stronger error than when violating one but confirming the other (i.e., “confusion”). In this study, regularity and intention seem to have provided enough precision when presented individually, such that presenting them together does not add certainty, but redundancy. Note that another major difference between the study of Pieszek et al. (2013) and this study consists in the very nature of the top-down expectations (visual vs. intention based), which might be implemented differently, producing effects at different latencies. Indeed, although in the study of Pieszek et al. (2013) the bottom-up versus top-down effects are being additive for the mean amplitudes around 105–130 msec, in this study we show that several components in the 100–200 msec latency ranges respond differently to bottom-up regularity versus top-down intention manipulations.

### **No N1 Effects for Top-Down Intention-based Modulations?**

We shall finally point out that the lack of N1 effects in the Intention condition is surprising, if we consider that previous studies bring forward the central contribution of motor intention to explain the N1 results typically found with self-generation paradigms (Timm et al., 2014; Hughes et al., 2013). Because we can clearly exclude the possibility that deviance has simply not been

detected in the Intention condition, based on the MMN and P3a effects, different explanations can be considered.

First, as already mentioned, it has been proposed that the N1 effects observed with self-generation studies represent “unspecific” responses following tones presented at rather long intervals (SanMiguel, Todd, et al., 2013) that do also not necessarily have to be contingent to the button presses (Horváth, 2013; Horváth, Maess, Baess, & Tóth, 2012). Additionally, even though regularity is not often considered as an explaining mechanism for the N1 effects reported with typical self-generation paradigms, it might play a role. Specifically, in contrast to self-generated tones where clear temporal regularity is established via self-pacing, clear temporal relations between consecutive tones cannot be as easily determined for externally generated tones (see Hughes et al., 2012, 2013, for similar arguments). In this study, the button presses determine sounds in all conditions at the same self-pacing rate—this means there are no temporal regularity differences between conditions and thus no effects are elicited at this level. Altogether, this leaves the later components around 200 msec as better candidates for the interpretation based around motor intention, idea previously suggested (SanMiguel, Todd, et al., 2013) and recently confirmed by LeBars and colleagues who showed that the N2b, but not the N1, is specific to intention-driven in comparison to stimulus-driven action (Le Bars et al., 2019).

Alternatively, a more trivial but also likely explanation is that the action-tone associations in the Intention condition were simply not strong enough to implement predictions at N1 sensory levels. In a comparable study, Hughes and colleagues asked participants to produce action-specific tones by inversely associating left versus right key presses with high versus low pitch tones (Hughes et al., 2013). In the predictable condition, these associations generated their corresponding tones with 100% probability, whereas in the unpredictable condition, two additional buttons for each hand generated high and low tones with equal chances. Comparing predictable versus unpredictable tones across both hands, they did not find an N1 effect—we thus replicate these findings. However, Hughes et al. (2013) found an N1 effect by looking at hand-specific prediction-congruent versus -incongruent tones. This result suggests that intention-based predictions at N1 levels can occur if a strong association is built-up between a hand-specific action and its corresponding tone. Correspondingly, it is possible that if the associations in our Intention condition were stronger, we would observe prediction errors at the N1 level. This hypothesis should be further tested—this is particularly worth investigating, because, if stronger action-tone associations in the Intention condition would lead to N1 effects (similar to the Regularity condition), this could in turn lead to additivity of bottom-up and top-down effects at N1 sensory levels.

### **Summary and Conclusion**

This study represents, to our best knowledge, the first direct comparison of intention and regularity effects, in

the context of action-related predictions. We show that the N1 component is modulated by tone regularity violations, but not by intention violations—one possible explanation is that it reflects a necessary consequence of neural adaptation. However, the MMN component is modulated both by regularity and intention violations. Because this effect is robust in all three conditions regardless of the global tone regularity (i.e., when controlled for neural adaptation), we argue that top-down predictions based on action intention are reflected at this sensory processing level. We did not observe stronger prediction errors when regularity and intention were concurrently violated (i.e., the two did not have additive effects). This suggests that, even though the two presumably have independent generators, converging predictions integrate (in contrast to prediction error responses adding up). Similar P3a effects across conditions point out that the deviance is further processed, regardless of whether the effects at the earlier sensory levels were based on regularity, intention, or both.

### Acknowledgments

We thank all participants who took part in this study. The data collection and analysis were conducted at the Cognitive and Biological Psychology Laboratory, University of Leipzig. This project was funded by the German Research Foundation ([www.dfg.de](http://www.dfg.de)) with project number GZ: SCHR 375/25-1. We would also like to thank Taavi Wenk for his help with the data collection as well as our lab colleagues for their useful comments regarding this study design.

Reprint requests should be sent to Betina Korka, Cognitive and Biological Psychology, Institute of Psychology, University of Leipzig, Neumarkt 9-19, D-04109 Leipzig, Germany, or via e-mail: [betina-christiana.korka@uni-leipzig.de](mailto:betina-christiana.korka@uni-leipzig.de).

### Notes

1. For the key-tone associations, in the Regularity condition, the first half of the participants generated a low tone and the second half generated a high tone as standards. In the Intention condition, half of the participants who generated the low standard tone in the Regularity condition generated a low standard tone with the left key and a high standard tone with the right key, and vice versa for the other half. The same associations were implemented for the second half of the participants who generated the high standard tone in the Regularity condition, thus far obtaining four quarters of possible associations (1: Regularity low-Intention left low-Intention right high; 2: Regularity low-Intention left high-Intention right low; 3: Regularity high-Intention left low-Intention right high; 4: Regularity high-Intention left high-Intention right low). In the Both condition, half of the participants (Quarters 1 and 4) pressed the left key frequently, and half of them pressed the right key frequently (Quarters 2 and 3). The key-tone associations were the same as in the Intention condition. Note that the standard tone for the frequently pressed key in the Both condition was the same as in the Regularity condition. For the condition order, permutations of Regularity, Intention, and Both conditions result in six different combinations, each of which was assigned to a participant within one quarter (6 condition orders  $\times$  4 quarters = 24 participants).

2. The measures computed by FASTER include the median slope of time course, slope of the power spectrum, spatial

kurtosis, Hurst exponent, and correlation with eye channels. ADJUST computes the spatial average difference, spatial eye difference, generic discontinuity spatial feature, maximum epoch variance, and temporal kurtosis. Finally, the measures computed by SASICA are low autocorrelation of time course, focal channel topography, focal trial activity, correlation with vertical EOG, and correlation with horizontal EOG. On average, 13.25 (range: 11–14) components per participant were removed.

### REFERENCES

- Bäß, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*, 137–143.
- Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, *83*, 120–131.
- Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 16.
- Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. *American Journal of EEG Technology*, *25*, 83–92.
- Chaumon, M., Bishop, D. V., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, *250*, 47–63.
- Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*, 732–744.
- 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*, 9–21.
- Dien, J. (2010). The ERP PCA toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, *187*, 138–145.
- Dien, J. (2012). Applying principal components analysis to event-related potentials: A tutorial. *Developmental Neuropsychology*, *37*, 497–517.
- Eaton, J. W., Bateman, D., Hauberg, S., & Wehbring, R. (2016). *GNU Octave* (Version 4.2.0.).
- Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 229–240.
- Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, *10*, 590–604.
- Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*, 215.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *360*, 815–836.
- Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*, 453–463.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23.
- Hari, R., Kaila, K., Katila, T., Tuomisto, T., & Varpula, T. (1982). Interstimulus interval dependence of the auditory vertex

- response and its magnetic counterpart: Implications for their neural generation. *Electroencephalography and Clinical Neurophysiology*, *54*, 561–569.
- Horváth, J. (2013). Action-sound coincidence-related attenuation of auditory ERPs is not modulated by affordance compatibility. *Biological Psychology*, *93*, 81–87.
- Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54–65.
- Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action-sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience*, *24*, 1919–1931.
- Horváth, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology*, *79*, 139–147.
- Hughes, G., Desantis, A., & Waszak, F. (2012). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, *139*, 133–151.
- Hughes, G., Desantis, A., & Waszak, F. (2013). Attenuation of auditory N1 results from identity-specific action-effect prediction. *European Journal of Neuroscience*, *37*, 1152–1158.
- Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences, U.S.A.*, *101*, 6809–6814.
- Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, *38*, 723–727.
- Jeffreys, H. (1961). *Theory of probability*. Oxford: Oxford University Press.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*, 1–16.
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*, 712–719.
- Kotchoubey, B., & Pavlov, Y. G. (2019). A Signature of passivity? An explorative study of the N3 event-related potential component in passive oddball tasks. *Frontiers in Neuroscience*, *13*, 365.
- Lange, K. (2013). The ups and downs of temporal orienting: A review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Frontiers in Human Neuroscience*, *7*, 263.
- Le Bars, S., Darriba, Á., & Waszak, F. (2019). Event-related brain potentials to self-triggered tones: Impact of action type and impulsivity traits. *Neuropsychologia*, *125*, 14–22.
- Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A: Optics, Image, Science, and Vision*, *20*, 1434–1448.
- Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge: Cambridge University Press.
- McCallum, W. C., & Curry, S. H. (1980). The form and distribution of auditory evoked potentials and CNVs when stimuli and responses are lateralized. *Progress in Brain Research*, *48*, 229–240.
- Mognon, A., Jovicich, J., & Bruzzone, L. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*, 229–240.
- Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, *42*, 25–32.
- 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*, 2544–2590.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425.
- Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, *48*, 162–175.
- Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods*, *192*, 152–162.
- Oprina, A., Popescu, A., Simion, E., & Simion, G. (2009). Walsh-Hadamard randomness test and new methods of test results integration. *Bulletin of the Transilvania University of Brasov. Mathematics, Informatics, Physics. Series III*, *2*, 93–105.
- Pieszek, M., Widmann, A., Gruber, T., & Schröger, E. (2013). The human brain maintains contradictory and redundant auditory sensory predictions. *PLoS One*, *8*, e53634.
- Prinz, W. (1997). Perception and action planning. *European Journal of Perception and Action Planning*, *9*, 129–154.
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374.
- SanMiguel, I., Saupe, K., & Schröger, E. (2013). I know what is missing here: Electrophysiological prediction error signals elicited by omissions of predicted “what” but not “when”. *Frontiers in Human Neuroscience*, *7*, 407.
- SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology*, *50*, 334–343.
- SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing silences: Human auditory processing relies on preactivation of sound-specific brain activity patterns. *Journal of Neuroscience*, *33*, 8633–8639.
- Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, *34*, 245–257.
- Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. *European Journal of Neuroscience*, *41*, 641–664.
- Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *NeuroReport*, *16*, 1781–1785.
- Timm, J., SanMiguel, I., Keil, J., Schröger, E., & Schönwiesner, M. (2014). Motor intention determines sensory attenuation of brain responses to self-initiated sounds. *Journal of Cognitive Neuroscience*, *26*, 1481–1489.
- Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, *77*, 65–72.
- Waszak, F., & Herwig, A. (2007). Effect anticipation modulates deviance processing in the brain. *Brain Research*, *1183*, 74–82.
- Widmann, A., Kujala, T., Tervaniemi, M., Kujala, A., & Schröger, E. (2004). From symbols to sounds: Visual symbolic information activates sound. *Psychophysiology*, *41*, 709–715.
- Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data—A practical approach. *Journal of Neuroscience Methods*, *250*, 34–46.
- Wolpaw, J. R., & Penry, J. K. (1975). A temporal component of the auditory evoked response. *Electroencephalography and Clinical Neurophysiology*, *39*, 609–620.