What exactly is missing here? The sensory processing of unpredictable omissions is modulated by the specificity of expected action-effects

Betina Korka1 | Erich Schröger1 | Andreas Widmann1,2

1Cognitive and Biological Psychology, Leipzig University, Leipzig, Germany
2Leibniz Institute for Neurobiology, Magdeburg, Germany

Correspondence
Betina Korka, Cognitive and Biological Psychology, Institute of Psychology, Leipzig University, Neumarkt 9-19, D-04109 Leipzig Germany.
Email: betina-christiana.korka@uni-leipzig.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: SCHR 375/25-1

Abstract
We select our actions according to the desired outcomes; for instance, piano players press certain keys to generate specific musical notes. It is well-described that the omission of a predicted action-effect may elicit prediction error signals in the brain, but what happens in the case of simultaneous effector-specific (by contrast to effector-unspecific) predictions? To answer this question, we asked participants to press left and right keys to generate tones A and B; based on the action-effect association, the tones’ identity was either predictable or unpredictable, while rarely, the expected input was omitted. Crucially, the data show that omissions following hand-specific associations reliably elicited a late omission N1 (oN1) component, by contrast to the hand-unspecific associations, where the late oN1 was rather weak. An additional condition where both key-presses generated a unique tone was implemented. Here, rare omissions of the expected tone generated both early and late oN1 responses, by contrast to the condition in which two simultaneous action-effect representations had to be maintained, where only late oN1 responses were elicited. Finally, omission P3 (oP3) responses were strongly elicited for all omission types without differences, indicating that a general expectation based on a tone presentation (rather than which tone), is likely indexed at this stage. The present results emphasize the top-down effects of action intention on the sensory processing of omissions, where unspecific (vs. specific) and multiple (vs. single) action-effect representations are associated with processing costs at the early sensory levels.

KEYWORDS
auditory processing, event-related potential, intentional action, predictive coding, principal component analysis

Abbreviations: %ERR, error percentages; EEG, electroencephalography; ERP, event-related potential; ICA, independent component analysis; P1T, predictable, 1 tone condition; P2T, predictable, two tones condition; PCA, principal component analysis; U2T, unpredictable, two tones condition.

Edited by: Sophie Molholm

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

© 2020 The Authors. European Journal of Neuroscience published by Federation of European Neuroscience Societies and John Wiley & Sons Ltd
INTRODUCTION

When acting, we know about the specific consequences of particular actions—for instance, professional pianists know which piano key to press, with which hand and specific finger, in order to produce unique tones. According to the ideomotor principle, once an association between a certain action type and a specific sensory consequence has been learnt, we select our actions in accordance with their sensory consequences in order to produce desired effects (Elsner & Hommel, 2001; Hommel, Müsseler, Aschersleben, & Prinz, 2001). Even though action-related sensory predictions have been extensively studied in the auditory system (for reviews, see Bendixen, SanMiguel, & Schröger, 2012; Horváth, 2015; Hughes, Desantis, & Waszak, 2013a), only a few studies looked at the specific role of top-down effects of action intention and action-effect predictions represented on a trial-by-trial basis. To this end, we further address the role of top-down modulations on action-effect predictions, by looking at the sensory processing of omissions in the context of effector-specific vs. effector-unspecific actions effects.

According to the predictive coding theory, brain predictions are implemented hierarchically via feedforward and feedbackward loops, where top-down information about the expected input travel from higher to lower cortical areas, and vice versa for the bottom-up information regarding the received input (Feldman & Friston, 2010; Friston, 2005). The goal is to minimize prediction error by constantly updating the generative model, in case the top-down expectations vs. the bottom-up sensory input do not correspond to each other. Regarding the action-effect predictions, a few studies have looked at the event-related potentials (ERPs) following tones that either matched or mismatched expectations that were based on previously learnt associations for left and right-hand key-presses. In this context, it was found that self-generated tones which were congruent (by contrast to incongruent) with previously learnt associations for left and right hands, lead to attenuation of the N1 (Hughes, Desantis, & Waszak, 2013b) and P3a ERP responses (Waszak & Herwig, 2007). These results point to a mechanism based on identity-specific action-effect predictions, rather than based on the mere coupling of tones with motor acts, as often postulated by the forward models of action control and action-related predictions (Knolle, Schröger, Baess, & Kotz, 2012; Martikainen, Kaneko, & Hari, 2005; Miall & Wolpert, 1996). In line with this, results of a recent study indicate that mismatching the action intention-based predictions elicit MMN and P3a responses, even when two tones that are inversely associated with left and right-hand key-presses are overall presented with equal chances. (Korka, Schröger, & Widmann, 2019).

In everyday situations, the predicted effects of our actions are occasionally absent (by contrast to different or mismatching)—for instance, pressing a button on the TV remote control might do nothing if the batteries are low, as opposed to changing the channel, as one would expect. It has been suggested that studying the responses to omitted, but highly expected stimuli allows uncovering the endogenous neural signature of predictions (Arnal & Giraud, 2012), by contrast to studying the responses to mismatching tones, which can only offer indirect evidence regarding the existence of a prediction (Schröger, Marzecová, & Sanmiguel, 2015). Specifically, in case of mismatching tones, the resulting prediction error confounds two aspects: first, the processing of input that has been delivered, but not predicted (i.e. the deviant/mismatching tone), and second, the lack of input that has been predicted, but not presented (Schröger et al., 2015). The study of omission responses thus allows disentangling a more “direct” measure of the prediction error. Following this reasoning, SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schröger (2013) showed that the brain responses to rare omissions of expected self-generated tones closely resembled the auditory ERP responses elicited by the self-generated tones themselves. Specifically, they asked participants to generate a click sound by pressing a key every 600–1,200 ms. In one condition, the sound was rarely omitted (in 12% of the cases), while in another condition, the sound was omitted in half of the trials. Omission N1 and N2 responses were only elicited for the rare omissions, where a relatively stable association between the action and its effect could be formed (SanMiguel, Widmann, et al., 2013). In another study, SanMiguel et al. further tested whether the omission responses uncover a prediction mechanism that is specific to the tone identity—they compared a condition where the key-press produced a unique (repetitive) output, with a condition where the key-press produced random tones, drawn on every trial from a pool of 48 environmental tones. Omission N1 (oN1), omission N2 (oN2) and omission P3 (oP3) responses were only elicited in the condition where the tone identity was unique and thus predictable (SanMiguel, Saupe, & Schröger, 2013). With enhanced statistical power and an adjusted EEG analysis pipeline, Dercksen et al. attempted to replicate the results of SanMiguel, Saupe, et al. (2013) and showed that oN1 and oP3 responses were also elicited in the random condition; nonetheless, these responses were indeed attenuated, by comparison to the condition in which tones had a unique identity (Dercksen, Widmann, Schröger, & Wetzel, 2020). To conclude, these results indicate that omissions represent a promising tool of studying action-related predictions and that top-down expectations regarding the precise stimulus identity seem to play an important role.

The present study takes a step forward and investigates whether omissions of effector-specific and effector-unspecific action-effects lead to differential sensory processing. Specifically, participants press left and right keys to generate tones A and B, while tone identity is either predictable
based on the action-effect association (left key-press elicits tone A, right key-press elicits tone B), or unpredictable (both key-presses generate both tones with equal probability). Contrasting with previous studies (Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013), here, the predictable vs. unpredictable conditions have the advantage of being physically identical in terms of incoming auditory stimulation. Thus, any differences between conditions can only be attributed to the influence of top-down expectations, as opposed to differences in stimulation. A third condition in which the tone identity is predictable but not effector-specific was implemented (participants pressed left and right keys to generate a unique tone), in order to help evaluate the processing costs of maintaining two vs. one sound representations. Finally, a silent key-presses condition provides a motor baseline. Importantly, this is physically identical to the omissions, but, in agreement with previous studies, we assume that here, the system should not be able to form any prediction (Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013; SanMiguel, Widmann, et al., 2013). With the motor signals across all conditions being presumably equal, we thus evaluate the differences between the omissions and motor wave as reflecting prediction-related activity.

First, in line with previous omission studies, we expect that a predictable tone identity based on the specific action-effects will lead to elicitation of the oN1, oN2 and oP3. Such omission responses should not occur or should be attenuated when the tone identity is effector-unspecific and thus unpredictable (Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013); however, we shall see if the system regards a binary (unspecific) outcome in the same way it regards an unspecific outcome referring to many more possibilities. Second, based on previous studies looking at mismatching action-effect predictions, we do not expect that simultaneous predictions should involve great processing costs compared to single tone predictions. However, to our knowledge, this represents the first study looking at the effects of multiple action-effect predictions on the processing of omissions; thus, we shall see if differences between maintaining two vs. one sound representations might be observable this way.

2.2 | Stimuli

Tones representing a dog bark, a bird chirp, a horn, pouring water and a telephone ring were chosen from the 48 environmental tones database collected by Wetzel, Widmann, & Schröger and used in previous studies in our laboratory (see e.g., Wetzel, Widmann, & Schröger, 2009, or SanMiguel, Saupe, et al., 2013). Each tone of 200-ms length including 10 ms rise-and-fall times and was presented binaurally through a pair of headphones (Sennheiser HD 25) at an intensity level of 65 dB SPL. A set of six descriptive icons (~\(1.9° \times 1.9°\) corresponding to either one of the five tones or to a silent key-press (i.e. an empty rectangle) were displayed to strengthen the expectations regarding the forthcoming stimulation and/or tone identity—see Figure 1a. These were presented in white on a black screen, left and right relative to the centre of the screen, on a 19" CRT monitor (G90fB, ViewSonic, resolution 1,024 × 768 pixels, refresh rate of 100 Hz) placed at a comfortable watching distance in front of the participant (~ 60 cm). Stimuli were delivered via the Psychotoolbox 3 (Kleiner et al., 2007) in combination with GNU Octave version 4.0.0 (Eaton, Bateman, Hauberg, & Wehbring, 2015), running on Linux OS.

2.3 | Apparatus and task

For the whole experiment duration, participants sat in a comfortable office chair in an electrically shielded, double-walled sound booth (Industrial Acoustics Company) and were required to fix their gaze on the fixation cross, presented in the centre of the screen. Their task was to press two keys, in turn, with their left and right index fingers, to generate tones (or silent presses in the Motor condition), according to the condition-specific key–tone associations. Figure 1a summarizes an example of key-press–tone associations for each of the experimental conditions. In the Predictable, two tones (P2T) condition, participants pressed the left key to generate a tone representing a bird chirp and the right key to generate a tone representing a dog bark. The tones were presented with a probability of 88%, while on the remaining 12%, omissions of the expected input occurred instead. In an Unpredictable, two tones (U2T) condition, participants pressed both the left and right keys to generate either the tone produced by a car horn or by pouring water. The two tones were presented with equal chances for both key-presses, thus making the forthcoming tone’s identity unpredictable. The tones were, like in the P2T condition, in agreement with the Declaration of Helsinki (code of approval: 2019.06.13_eb_15). Participants received either compensation of 8 euros/hour or course credits.

2 | MATERIALS AND METHODS

2.1 | Participants

Data were collected from 30 participants (13 male, mean age = 25.1 years, age range: 19–34), all of which 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, nor was taking any prescribed drugs. The Ethics Advisory Board of Leipzig University approved the study procedure, according to the condition-specific key–tone associations.
presented with 88% chance, while on the remaining 12% of the trials, omissions occurred instead. In a Predictable, one tone (PIT) condition, participants’ task was to press both keys to generate a telephone ring in 88% of the cases, while in remaining 12% of the trials, omissions occurred instead. Finally, in a Motor condition, participants performed silent key-presses, which, importantly, are physically identical with the omission trials, but where, in accordance with previous studies (Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013; SanMiguel, Widmann, et al., 2013), we assume no prediction-related activity can occur.

Note that the described key-press—tone associations represent only one possible scenario only (i.e. they have been counterbalanced between participants). The omissions in all conditions (just like in the silent presses) were defined as tones of intensity 0. Participants’ instructions were to press a key of their choice every second, without producing fixed sequences of left-right key-presses and equally often throughout one block. They were made aware that rarely, the expected tones will not be presented—we wanted to avoid the possibility that participants stop throughout the block and report that something “wrong” happened. The two keys had dimensions of 3.8 × 4.5 cm and were placed on the desk in front of the participant. Importantly for the scope of this study, we used custom-built infrared photoelectric sensor-based keys, which, by contrast to membrane or mechanical keys have the advantage of being completely silent, while still providing tactile feedback associated with the press.
2.4 Experimental procedure

One experimental session lasted approximately one hour in which a total number of 16 experimental blocks were recorded. Each of the omission conditions (P2T, U2T and P1T) consisted of five blocks with a duration of about three minutes each. The Motor condition consisted of one single block with the duration of about two minutes. Four shorter practice blocks were run in the beginning of each condition, while blocks corresponding to the same condition were run one after another. In between blocks, participants could take self-paced breaks. The condition order and the associations between key-presses and tones for the three-tone conditions were counterbalanced between participants, as follows. For the condition order, full permutations corresponding to the four conditions results in 24 combinations; in order to reach 30 (i.e. our number of participants), 6 out of the 24 responding to the four conditions results in 24 combinations; in participants, as follows. For the condition order, full permutations corresponding to the four conditions results in 24 combinations; in order to reach 30 (i.e. our number of participants), 6 out of the 24 combinations were repeated with the constraint that each condition appeared at each position (1st, 2nd, 3rd or 4th) at least seven times. For the key-presses–tones associations, full permutations corresponding to the five tones results in 120 possible combinations, however, given that in the U2T condition the associations for the left and right key-presses are irrelevant (i.e. both hands generate both tones equally likely), the number of permutations was reduced to 60. We used the Latin squares method to further reduce the number of combinations to 30, such that each of the five tones was associated six times with the P1T condition, six times with the left key-press in P2T and six times with the right key-press in the P2T condition. For the U2T condition, every possible pair of tones (10, in total) appeared three times.

Figure 1b illustrates an example of trial progression. The fixation cross along with the icons indicating the expected tones for the left and right key-presses were present on-screen at all times. The screen feedback indicating the left and right key-presses to be performed was presented in numbers and percentages at the left and right sides of the fixation cross (on top of the icons), while the time interval between two consecutive key-presses was presented starting from the third trial underneath the fixation cross—this information was updated 500 ms after every key-press. That is, following every press, these numbers and percentages adjusted accordingly—this helped ensure that participants pressed the two keys equally often and at the indicated pace of about one second. If a participant’s pace was much faster or slower than one second by ±500 ms, a corresponding error message (“Too short/Too long”) was displayed on the screen, instead of the timing between the key-presses. These were considered timing errors, and the percentage of trials containing them was analysed offline. All the on-screen information was presented in white on a black background. An algorithm detecting stereotypic repetitions (multiple of 2, 3 or 4) associated with the codes for the left and right key-presses was programmed within the experiment. The sequence of key-presses was thus monitored online by the experimenter, and if too many stereotypical patterns were detected, the participants were verbally warned at the end of the block. Note that the mere purpose of this was to ensure that participants did not produce an excessive amount of fixed order sequences of left and right key-presses. The tone onset immediately followed the key-press (with a delay of max. 5 ms). Each of the five omission blocks in the P2T, U2T and P1T conditions consisted of 176 tones and of 24 omissions, representing 88% and 12% of the trials in that block, respectively. In the P2T condition, left key-presses generated tone A (in Figure 1a, a bird chirp) and the right key-presses generated tone B (in Figure 1a, a dog bark) on 44% of the trials. In the U2T condition, tones C and D (in Figure 1a, a horn noise and water pouring) were generated equally likely by both left and right key-presses, each on 44% of the trials. In the P1T condition, each of the left and right key-presses generated tone E (in Figure 1a, a telephone ring) on 44% of the trials. Omissions in all conditions were generated with 6% chance for each of the left and right key-presses.

In total, 880 tones and 120 omissions were recorded for each of the P2T, U2T and P1T conditions. The Motor block consisted of 120 silent key-presses, 60 for each of the left and right keys. For the P2T, U2T and P1T conditions, the practice blocks consisted of 50 trials, 25 tones corresponding to each of the left and right key-presses. Note that in the practice blocks, no omissions but only tones were presented, in order to help build-up specific action-effect associations in the P2T and P1T conditions (or the lack of clear associations in U2T condition). The order of tones and omissions was randomized, with the constraint that there were no consecutive omissions and the first two trials of every block were always tones.

2.5 EEG data recording

EEG data were continuously recorded at a sampling rate of 500 Hz with a system equipped with 32 Ag-AgCl active electrodes, using a BrainAmp amplifier and the Vision Recorder software (Brain Products™ GmbH). Twenty-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, while three electrodes were used to record EOG activity, two of which were placed on the left and right outer canthi and one below the left eye.

2.6 EEG preprocessing

EEGLAB MATLAB-based software (Delorme & Makeig, 2004) was used for the preprocessing. Data were filtered using a 0.2 Hz high-pass and 45 Hz low-pass windowed sinc finite
impulse response (FIR) filter (Hamming window, filter order 8250—high-pass and 166—low-pass), in accordance with the recommendations of Widmann, Schröger, & Maess (2015). On average, 0.65 channels (range: 0–4) containing extreme amplitudes were removed using a deviation criterion (threshold = 3) which “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 500 ms). Epochs with amplitudes exceeding a 500 μV amplitude difference threshold were removed. An independent component analysis (ICA) was computed on the raw data, which were first filtered with a 1 Hz high-pass and 45 Hz low-pass filter in order to optimize the ICA decomposition, epoched (~200 to 500 ms relative to tone presentation) and cleaned by removing the same bad channels and epochs detected at the earlier step. The obtained weights were stored and transferred to the 0.2 Hz high-pass filter datasets. The removal of components containing eye-related and muscle artefacts was done based on visual inspection and paired with the recommendations computed by SASICA, which refer to low auto-correlation of time course, focal channel topography, focal trial activity, correlation with vertical EOG and correlation with horizontal EOG. (Chaumon, Bishop, & Busch, 2015). On average, 7.48 components per participant (range: 6–9) were removed. The missing channels were interpolated using the built-in EEGLAB spherical interpolation function. Data were baseline corrected using the first 100 ms of the pre-stimulus interval—we did not use the whole pre-stimulus (200 ms) interval in order to avoid potentially introducing motor-related signals into the post-stimulus activity. Epochs with amplitudes still exceeding a 200 μV Delta threshold after the ICA corrections were removed—overall, removed epochs (summing both the 500 and 200 μV Delta thresholds) represented on average ~ 0.67% from the total number of trials (range: 0% - 4.83%). Finally, condition-specific averages were calculated.

2.7 | PCA analysis

A temporal principal component analysis (PCA) was performed using the ERP PCA toolkit MATLAB-based toolbox (Dien, 2010). We computed the PCA on the individual averages corresponding to the three omission types (P2T, U2T and U1T conditions) and Motor condition, using a Geomin rotation with a covariance relationship matrix and no weighting (Scharf & Nestler, 2019). Horn’s parallel test was used to determine the number of components to be retained.

2.8 | Statistical analysis

The analysis focuses on the three omission types and motor wave. Each component of interest identified by temporal PCA was separately analysed using Bayesian paired samples t tests and additionally, their complementary frequentist paired samples t tests. Each of the three omission types was directly contrasted with the motor wave, to test for prediction-related activity (P1T vs. Motor; P2T vs. Motor; U2T vs. Motor). The omissions were then directly contrasted as follows: P2T vs. U2T in order to test if effector-specific vs. effector-unspecific action-effects involve any sensory processing benefits/costs, and P2T vs. P1T, to help evaluate the costs of maintaining two vs. one sound representations. The analysis was conducted using the JASP 0.9.1.0 software. The Bayes factor (BF10) was calculated using 10,000 Monte-Carlo sampling iterations; the null hypothesis corresponded to a standardized effect size δ = 0, while the alternative hypothesis was defined as a Cauchy prior distribution centred around 0 with a scaling factor of r = 0.707 (Rouder, Morey, Speckman, & Province, 2012). In line with the Bayes factor interpretation (Jeffreys, 1961; Lee & Wagenmakers, 2013) and with previous studies reporting Bayes factors (Korka et al., 2019; Marzecová et al., 2018; Stuckenberg, Schröger, & Widmann, 2019), data were taken as moderate evidence for the alternative (or null) hypothesis if the BF10 was greater than 3 (or lower than 0.33), while values close to 1 were considered only weakly informative. Values greater than 10 (or smaller than 0.1) were considered strong evidence for the alternative (or null) hypothesis. For the frequentist complementary analysis, statistical significance was defined at the .05 alpha level, and results are reported including Cohen’s d effect size (d).

3 | RESULTS

3.1 | Timing errors

Timing errors (see the Experimental procedure section) were calculated as error percentages (%ERR) relative to the total number of omissions/silent presses in each condition—trials containing them were excluded from the further ERP analyses. On average, participants made 0.21 %ERR in P2T (SD = 0.45%, range = 0%–2%), 0.12 %ERR in U2T (SD = 0.24%, range = 0%–1%), 0.13 %ERR in P1T (SD = 0.24, range = 0%–1%) and 2.13 %ERR (SD = 5.84%, range = 0%–28.33%) in the Motor condition. Based on the normalized z-scores for each participant and condition, we excluded one participant from the further analyses (including the ERP analysis), whose error rate was more than 4 standard deviations above the mean in the Motor condition (i.e. 28.33% relative to the 2.13% mean). Condition-specific error rates from the remaining 29 participants were included in a one-way Bayesian rANOVA (P2T, U2T, P1T, Motor), which provided evidence for condition differences (BF10 = 4.32 ± 0.67%). However, post hoc comparisons were uninformative regarding the differences between the Motor condition and P2T.
(BF$_{10}$ = 0.84 ± <0.001%), U2T (BF$_{10}$ = 1.18 ± <0.001%) and P1T (BF$_{10}$ = 1.01 ± <0.001%) conditions, but they did support the null hypothesis regarding the differences between P2T and U2T (BF$_{10}$ = 0.25 ± 0.003%), P2T and P1T (BF$_{10}$ = 0.23 ± 0.001%) and P1T and U2T (BF$_{10}$ = 0.2 ± <0.001%). In sum, the overall small error rates (less than 1.3% in all conditions after exclusion of one participant) suggest that the task instructions were closely followed and that participants pressed the two buttons at the suggested pace.

### 3.2 ERP PCA results

Figure 2 displays the grand-average ERPs along with the PCA results. Thirteen principal components explaining over 95% of the epoch variability were retained, as suggested by Horn’s parallel test (see Figure 2b, last row). Two regions of interest composed of frontocentral (Fz, Cz) and temporal (T7, T8) electrodes, respectively, display the grand-average ERPs for all omissions types and the motor wave, by comparison to the reconstructed PCA waves obtained by the sum of the 13 retained components. Note that the PCA reconstruction waves correspond well to the grand-average waves, while some of the noise present in the grand-average waves has additionally been removed. Out of the 13 retained components, we further analysed four, presumably representing early and late oN1 and early and late oP3 responses, respectively—the selection of these components of interest was based on latency and topographical information.

Note that the principal components are ordered not by chronological peak latency, but by the explained variance. Accordingly, the early and late oP3 components peaking at

**FIGURE 2** ERP PCA results. Column (a): Grand-average ERPs display the P2T, U2T, P1T and Motor waves, for an average of the frontocentral Fz and Cz electrodes (top), and an average of temporal T7 and T8 electrodes (bottom). Column (b): Following the PCA analysis, 13 components explaining more than 95% of the whole epoch variability were retained. The waves representing the sum of these 13 components for the same frontocentral (top) and temporal (middle) electrodes displaying the P2T, U2T, P1T and Motor conditions correspond well to the grand-average ERPs, while additionally removing some noise. The 13 retained components are presented individually (bottom): out of these, four components (represented in darker lines) presumably representing early and late oN1 and oP3 responses were further analysed. ERP, event-related potential; PCA, principal component analysis [Colour figure can be viewed at wileyonlinelibrary.com]
256 and 334 ms together explain about 68.6% of the variability, while the early and late oN1 components peaking at 94 and 152 ms together explain about 2.3% of the variability. A motor component peaking around time 0 (i.e. around the button-press) explains about 3.4% of the variability, while the late epoch activity is represented by a component peaking at 494 ms and explaining about 14.6% of the variability. A P2-like component peaking around 200 ms further explains about 1.6% of the variability—we did however not analyse this (despite being in a latency of interest), as no specific hypothesis was formed for an omission-related P2. Finally, the remaining six principal components explain less than 1% of the variability, each. The time-variant loadings of the components reflect their contribution to the voltage maps at each point in time (see Figure 2b, last row). The time-invariant component scores represent the contribution of each component to the ERP wave—these have been further analysed statistically, for the oN1 and oP3 components of interest, which are next reported. For a summary of the Bayesian and frequentist statistical results, see Table 1.

3.3 Early and late oN1

Figure 3 illustrates the early and late oN1 ERPs and corresponding topographical voltage maps, along with the distribution of scores for every condition, after subtraction of the motor wave. The oN1 (both early and late) has a temporal bilateral distribution peaking at electrodes T7 and T8—the analysis therefore focused on an average of these two. For the early oN1, concerning the existence of prediction-related activity (i.e. differences between each of the three omission types and the Motor wave), the Bayesian t tests provided strong support the alternative hypothesis in the P1T (P1T vs. Motor: BF$_{10}$ = 177.33), but not in the P2T (P2T vs. Motor: BF$_{10}$ = 0.65), nor in the U2T (U2T vs. Motor: BF$_{10}$ = 0.59) conditions. Similarly, the frequentist t tests indicated a significant effect in the P1T condition (P1T vs. Motor: t(28) = −4.37, p = <.001, d = −0.812), but not in the P2T (P2T vs. Motor: t(28) = −1.64, p = .110, d = −0.306), nor in the U2T conditions (U2T vs. Motor: t(28) = −1.57, p = .126, d = −0.293). Regarding the differences between effector-specific vs. effector-unspecific action-effects, the Bayesian t tests supported the null hypothesis (P2T vs. U2T: BF$_{10}$ = 0.24), while the evidence is rather inconclusive in the case of unique vs. simultaneous action-effect representations (P2T vs. P1T: BF$_{10}$ = 0.85). Correspondingly, the frequentist t tests yielded non-significant effects for both comparisons (P2T vs. U2T: t(28) = −0.69, p = .495, d = −0.128; P2T vs. P1T: t(28) = 1.83, p = .078, d = 0.340).

For the late oN1, the Bayesian comparisons against the motor wave brought strong support for the existence of prediction-related activity in the P2T (P2T vs. Motor: BF$_{10}$ = 15.14) and P1T (P1T vs. Motor: BF$_{10}$ = 21.04) conditions, while in the U2T condition, the evidence was rather weak (U2T vs. Motor: BF$_{10}$ = 1.87). The complementary

| Comp. | Prediction-related activity | Condition differences |
|-------|-----------------------------|-----------------------|
|       | Contrast | BF$_{10}$ | t   | p    | d    | Contrast | BF$_{10}$ | t   | p    | d    |
| Early oN1 | P2T vs. Motor | 0.65 | −1.64 | .110 | −0.306 | P2T vs. U2T | 0.24 | −0.69 | .495 | −0.128 |
|          | U2T vs. Motor | 0.59 | −1.57 | .126 | −0.293 | P2T vs. P1T | 0.23 | 0.18 | 0.856 | 0.034 |
|          | P1T vs. Motor | 177.33 | −4.37 | <.001 | −0.812 | P2T vs. P1T | 0.85 | 1.83 | .078 | 0.340 |
| Late oN1 | P2T vs. Motor | 15.14 | −3.32 | .002 | −0.618 | P2T vs. U2T | 1.50 | −2.17 | .038 | −0.404 |
|          | U2T vs. Motor | 1.87 | −2.30 | .029 | −0.427 | P2T vs. P1T | 0.23 | 0.60 | .549 | 0.113 |
|          | P1T vs. Motor | 21.04 | −3.47 | .002 | −0.645 | P2T vs. P1T | 0.22 | −0.51 | .613 | −0.095 |
| Early oP3 | P2T vs. Motor | 12.21 | 3.22 | .003 | 0.600 | P2T vs. U2T | 0.22 | −0.51 | .613 | −0.095 |
|          | U2T vs. Motor | 23.34 | 3.51 | .002 | 0.653 | P2T vs. P1T | 0.20 | 0.18 | .856 | 0.034 |
|          | P1T vs. Motor | 14.83 | 3.31 | .003 | 0.616 | P2T vs. P1T | 0.22 | 0.18 | .856 | 0.034 |
| Late oP3 | P2T vs. Motor | 2,649.6 | 5.46 | <.001 | 1.014 | P2T vs. U2T | 0.22 | 0.48 | .634 | 0.089 |
|          | U2T vs. Motor | 233.1 | 4.48 | <.001 | 0.833 | P2T vs. P1T | 0.23 | −0.58 | .566 | −0.108 |
|          | P1T vs. Motor | 1,207.8 | 5.14 | <.001 | 0.956 | P2T vs. U2T | 0.22 | 0.48 | .634 | 0.089 |

For each of the early and late oN1 and oP3 components, all omission types are tested against the motor wave to identify the prediction-related activity (left). For each of the four components, the omissions in the P2T and U2T conditions are directly compared to determine if effector-specific vs. effector-unspecific action-effects cause differential prediction error signals, while the direct comparison of omissions in the P2T vs. P1T conditions helps evaluate the costs of maintaining two vs. one sound representations (right). For all comparisons, Bayes factors (for which all corresponding errors ± err% ≤ 0.002) are reported along with the complementary frequentist results (t-values, p-values and Cohen’s d effect sizes). Contrasts representing strong Bayesian support for the alternative hypothesis, as well as significant differences in the frequentist analysis, are highlighted in bold. Brackets mark that none of the two conditions directly compared elicited prior prediction-related activity (by comparison to the motor wave).
frequentist $t$ tests indicated significant differences between each of the three omission types and the Motor wave (P2T vs. Motor: $t(28) = -3.32, p = .002, d = -0.618$; P1T vs. Motor: $t(28) = -3.47, p = .002, d = -0.645$; U2T vs. Motor: $t(28) = -2.30, p = .029, d = -0.427$). Regarding the differences between effector-specific vs. effector-unspecific action-effects, the Bayesian $t$ tests offered only weak evidence for the alternative hypothesis (P2T vs. U2T: BF$_{10} = 1.50$), while the null hypothesis was supported for the differences between the unique vs. simultaneous action-effect representations (P2T vs. P1T: BF$_{10} = 0.23$). Comparably, the frequentist $t$ tests indicated a significant effect regarding the effector-specific vs. effector-unspecific differences (P2T vs. U2T: $t(28) = -2.17, p = .038, d = -0.404$), and a non-significant effect regarding the unique vs. simultaneous representations (P2T vs. P1T: $t(28) = 0.60, p = .549, d = 0.113$). In sum, these data provide reliable evidence for early and late oN1 elicitation following omissions in the P1T condition. Further, a late oN1 effect has been reliably elicited for the omissions in P2T condition, while the evidence is less convincing regarding the omissions in the U2T condition. Finally, these data provide (rather weak) evidence for effector-specific versus effector-unspecific differences (P2T vs. U2T conditions), at the late oN1 level.

### 3.4 Early and late oP3

Figure 4 illustrates the early and late oP3 ERPs and corresponding topographical voltage maps, along with the distribution of scores for every condition, after subtraction of the motor wave. The early oP3 has a frontocentral distribution—we thus included in the analysis an average of electrodes Fz, FC1, FC2 and Cz. Concerning the existence of
prediction-related activity (i.e. differences between each of the three omission types and the Motor wave), the Bayesian $t$ tests brought strong evidence in favour of the alternative hypothesis for all three omission types (P2T vs. Motor: $BF_{10} = 12.21$; U2T vs. Motor: $BF_{10} = 23.34$; P1T vs. Motor: $BF_{10} = 14.83$). Similarly, the frequentist comparisons indicated significant effects in all three conditions (P2T vs. Motor: $t(28) = 3.22$, $p = .003$, $d = 0.600$; U2T vs. Motor: $t(28) = 3.51$, $p = .002$, $d = 0.653$; P1T vs. Motor: $t(28) = 3.31$, $p = .003$, $d = 0.616$). Regarding the differences between effector-specific vs. effector-unspecific action-effects, as well as between unique vs. simultaneous action-effect representations, the Bayesian $t$ tests supported the null hypotheses (P2T vs. U2T: $BF_{10} = 0.22$; P2T vs. P1T: $BF_{10} = 0.20$). The frequentist effects concurred, indicating non-significant differences for both comparisons (P2T vs. U2T: $t(28) = −0.51$, $p = .613$, $d = −0.095$; P2T vs. P1T: $t(28) = 0.18$, $p = .856$, $d = 0.034$).

The late oP3 is characterized by a widely distributed central activation pattern—we thus analysed a region of interest consisting of the FC1, FC2, Cz, CP1 and CP2 electrodes. The late oP3 results mirror the early oP3 results: that is, the Bayesian $t$ tests strongly supported the existence of differences between all three omission types and the Motor wave (P2T vs. Motor: $BF_{10} = 2,649.6$; U2T vs. Motor: $BF_{10} = 233.1$; P1T vs. Motor: $BF_{10} = 1,207.8$). Likewise, all three frequentist comparisons yielded significant differences (P2T vs. Motor: $t(28) = 5.46$, $p = <.001$, $d = 1.014$; U2T vs. Motor: $t(28) = 4.48$, $p = <.001$, $d = 0.833$; P1T vs. Motor: $t(28) = 5.14$, $p = <.001$, $d = 0.956$). Regarding the differences between effector-specific vs. effector-unspecific action-effects, as well as between unique vs.

---

**FIGURE 4** Early and late oP3 components. Column (a): The early oP3 peaks at 256 ms and is largest over a region of interest (ROI) composed of Fz, FC1, FC2 and Cz electrodes. Column (b): The late oP3 peaks at 334 ms and is largest over a ROI composed of FC1, FC2, Cz, CP1 and CP2 electrodes. The motor-corrected waves for both components display a mean of the electrodes representing the component-specific ROIs for the omission-evoked responses in the P2T, U2T and P1T conditions (top row). For all three omission types, the component-specific waves (non-faded lines) are displayed along with the “reconstruction waves” representing the sum of the 13 retained components (faded lines). The condition-specific topographical maps for both components are calculated based on spherical spline interpolation and illustrate the omissions–motor differences of activation (middle row). For both components, the violin plots display the condition-specific omissions–motor component scores, where the estimated density distributions contain the individual scores indicated by the smaller dots, and the means indicated by bigger black dots (bottom row) [Colour figure can be viewed at wileyonlinelibrary.com]
simultaneous action-effect representations, the Bayesian tests once again supported the null hypotheses (P2T vs. U2T: BF\textsubscript{10} = 0.22; P2T vs. P1T: BF\textsubscript{10} = 0.23). Similarly, both comparisons yielded non-significant frequentist effects (P2T vs. U2T: t(28) = 0.48, p = .634, d = 0.089; P2T vs. P1T: t(28) = −0.58, p = .566, d = −0.108). In sum, these data provide strong evidence for the elicitation of early and late oP3 responses following omission in all three conditions, with no condition differences.

4 | DISCUSSION

After a certain amount of practice, a piano player will know which keys to press, and in which order, to produce a melody—the piano player, and similarly, any agent acting in order to produce desired effects in the environment, are able to predict the sensory consequences of their own actions and select one action course or another accordingly. It is well-described that the omission of a predicted action-effect may elicit prediction error signals in the brain (Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013; SanMiguel, Widmann, et al., 2013); however, our ability to represent simultaneous predictions, in particular, effector-specific vs. unspecific action-effect predictions represented on a trial-by-trial basis, has been barely explored. To address this, we asked participants to press left and right keys to produce unique vs. two, and hand-specific vs. unspecific tones, which were rarely omitted. We aimed to answer two main questions: first, do hand-specific vs. hand-unspecific key–tone associations and second, does maintaining one vs. two sound representations involve any processing costs/benefits modulating the sensory processing of omissions? The results indicate that indeed unpredictable hand-unspecific (by contrast to predictable, hand-specific) associations involve some costs reflected in the oN1 response; the oN1 processing level also reflects costs associated with representing two (by contrast to one) sounds. Further, oP3 responses were strongly elicited for all omission types, indicating that there seem to be no costs or benefits associated with maintaining specific vs. unspecific, or single vs. simultaneous representations at this later processing level. Finally, no oN2 was found, by comparison to previous action-effect omission results (Dercksen et al., 2020; SanMiguel Saupe, & Schröger, 2013; SanMiguel, Widmann, et al., 2013). We next discuss these findings in more detail.

4.1 Costs of maintaining effector-specific vs. effector-unspecific sound representations at early processing levels

In line with our expectations and with previous results, the present data indicate that being able to determine the specific identity of the forthcoming tone plays a role for predictions at early auditory processing levels. That is, we show that when two different tones are inversely generated by left and right key-presses (i.e. are predictable based on the specific action-effect associations), late oN1 responses are reliably elicited in the P2T condition, by comparison to the (physically identical) motor wave. By contrast, when the two tones are generated with equal likelihoods by both key-presses in the U2T condition (i.e. are unpredictable due to the unspecific action-effect associations), the evidence regarding the existence of a late oN1 response by comparison to the motor wave is rather weak. Note that the motor wave comparisons as means of determining the existence of reliable omission-related activity represent a commonly used approach (see also Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013; SanMiguel, Widmann, et al., 2013) whose validity can be justified as follows. Regardless of the potential differences across experiments (the amplitude and morphology of the motor wave might vary depending on the type of keys used or required strength), within the same experiment, we can assume that all motor potentials are equal across conditions if participants press the same keys. Thus, if extra activation is observed in the omission trials by comparison to the motor wave, it can plausibly be related to sound expectation. When further and directly comparing the late oN1 responses between the P2T vs. U2T omission types, there also seem to be some rather weak differences. Nonetheless, the Bayes Factors regarding the motor wave comparisons indicate the existence of costs associated with maintaining effector-unspecific predictions in the U2T condition (BF\textsubscript{10} = 1.87), relative to the effector-specific associations in the P2T condition (BF\textsubscript{10} = 15.14).

SanMiguel et al. showed that omission responses were only elicited if first, a stable action-effect contingency is formed—that is, the action generates the expected effect more often than it generates an omission (SanMiguel, Widmann, et al., 2013). Second, when the tone identity is predictable, by contrast to randomly selected on every trial out of multiple possibilities, the magnitude of the omission responses is further modulated (Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013). Our results extend this and show how omission responses to predictable action-affects are also elicited when controlling for the differences in physical stimulation between the predictable and unpredictable conditions. Namely, by contrast to SanMiguel, Saupe, et al. (2013) and Dercksen et al. (2020), here, the hand-specific P2T and hand-unspecific U2T conditions were identical in terms of physical stimulation, meaning that the two tones were triggered overall with equal chances, the only difference consisting in the hand-tone associations. Thus, this demonstrates that the condition differences that we report can only be attributed to the influence of top-down expectations, as opposed to differences in stimulation and/
or differences regarding the neural refractoriness states corresponding to tones of different frequencies (Jacobsen & Schröger, 2001).

The top-down expectations represent a crucial element of the predictive coding theory (Friston, 2005). Specifically, the brain’s processing hierarchy involves bottom-up information (incoming stimuli) and top-down expectations which are constantly contrasted and integrated across multiple levels in the cortical feedforward/feedback loops (Lee & Mumford, 2003). Note that at a computational level, one subtlety of the predictive coding theory is that the top-down prediction units do not directly cause “expectations” in the sensory units, but a second-order stream is involved, where the precision gain defined as “the best estimate of the reliability or inverse variance of prediction errors” is modulated by prior knowledge (Friston, 2018). Nevertheless, experience-dependent expectations can be derived from multiple sources, such as overall probability or other regularities, multimodal associations (i.e. visual–auditory or auditory–visual), or intention-based action-effect couplings. Many action-effect studies failed to distinguish between the effects of overall probability vs. action intention. That is, even though the role of action intention as a main source of information driving predictions has been previously discussed in the literature (Hughes et al., 2013b; Waszak & Herwig, 2007), only recently did a study point out that the its effects on early sensory processing can be observed without global regularity of the standard tone (Korka et al., 2019). From this perspective, the present results are especially illustrative; here, the omission-related activity likely represents the endogenous neural signature of prediction (Arnal & Giraud, 2012) that appears to be exclusively based on intention and specific action-effect couplings.

Note that by discussing the differences between the effector-specific vs. effector-unspecific oN1 results as a function of top-down driven predictions, we argue that there are processing costs associated with an unpredictable tone identity. We do not argue that the hand-unspecific predictions do not exist or cannot be accounted for by prior expectations. In fact, the omission results of Dercksen et al. (2020) clearly indicate that prediction-related activity is also elicited when the tone identity cannot be determined (Dercksen et al., 2020). Indeed, here, the Bayesian evidence does not suggest the absence of an oN1 effect in the U2T condition, but only that the effect is less reliable. In fact, it seems likely that here, as opposed to learning the absence of a stable association, participants rather learnt the existence of a binary association. Additionally, as we discuss later based on the oP3 results, it is also possible that even at this early processing level, participants form somewhat general predictions based on a tone occurring, rather than based on which specific tone will occur.

### 4.2 Costs of maintaining one vs. two sound representations at early processing levels

Our data shows that it does make a difference whether the system has one or two sounds to represent. Specifically, the motor wave comparisons indicate that when both actions generate the same tone in the PIT condition, the oN1 response is already strong around 94 ms, continuing into the later component around 152 ms. However, when each of the two actions generates specific tones in the P2T condition, only the later oN1 component is convincingly elicited. Even though the results provided by the direct comparison of the two omission types for the early oN1 component are rather inconclusive, the evidence provided by the motor comparisons makes a compelling argument for the costs associated with maintaining two vs. one sound representations ($BF_{10} = 0.65$ vs. $BF_{10} = 177.33$).

The system’s ability to maintain multiple vs. single representations might be influenced by the allocation of attentional resources that are known to have effects on sensory processing. The consensus is that attended stimuli are preferentially processed (Schröger et al., 2015). Conversely, in the case of unattended stimuli, the auditory evoked responses are weaker (Hillyard, Hink, Schwent, & Picton, 1973). Here, the very low timing error rates (see Timing Errors in the Results section) indicate that participants invested considerable effort into performing the task well; they were additionally required to monitor the generated sequences in order to avoid fixed patterns and an unequal number of left-right key-presses. Thus, it is likely that resources were allocated to performing the task itself, rather than actively attending to the sounds (and omissions). In a relevant study, Jones, Hughes, and Waszak (2013) looked at the orthogonal effects (i.e. referring to different stimulus features) of action-effect predictions and attention on the processing of tones, using EEG. A prediction $\times$ attention interaction effect was found only relatively late, between 200 and 300 ms, while at the earlier N1 level, only a typical attention effect was found, that is increased activation for all tones in the attended by contrast to unattended condition (Jones et al., 2013). In line with this, it might be that here, the overall (low amplitude) oN1 effects across conditions are explained by low attentional resources for the processing of tones/omissions. Still, why should there be costs associated with maintaining two predictions? It has been shown before that the system is able to maintain simultaneous (Horváth, Czigler, Sussman, & Winkler, 2001) and even contradictory predictions (Dürschmid et al., 2019; Pieszek, Widmann, Gruber, & Schröger, 2013), the ability to maintain multiple auditory stimulus configurations having further been tested in computational models (Mill, Böhm, Bendixen, Winkler, & Denham, 2011). While some costs between representing two vs. one prediction might indeed exist, they likely become more evident in the context of a rather demanding
 task. Indeed, previous action-effect studies looking at the additive (as opposed to interactive) effects of attention and prediction on the N1 component suggest that the two can have joint effects at this level (Saupe, Widmann, Trujillo-Barreto, & Schröger, 2013; Timm, SanMiguel, Saupe, & Schröger, 2013). Yet, it is for future research to establish whether this is also true in the case of omission-related activity.

Alternatively, the recently proposed model of the opposing process theory postulates that optimal perception is achieved by an initial prediction mechanism that preactivates the sensory units towards expected events (Press & Yon, 2019), followed by a later prediction error mechanism that enhances the sensory activity following “surprising” events (Press, Kok, & Yon, 2020). According to the initial prediction mechanism and in line with recent fMRI work, sharpening during action occurs because predicted signals are enhanced and unpredicted signals are suppressed (Yon, Gilbert, de Lange, & Press, 2018). According to the later prediction error mechanism, the magnitude of the response following surprising outcomes is proportional to the strength and precision of the initial prediction. On the one hand, in the context of our results, the initial prediction mechanism potentially implies that while tone A is expected, the units tuned to tone A will be activated and other units will be suppressed—this would presumably elicit the early oN1 effect in the P1T condition, while a single tone was expected. However, collapsing signals where both tones A and B were expected on different trials depending on the choice of action, mutually competitive signals cancel each other out—this would then explain the lack of an early oN1 in the P2T condition. The second mechanism of this model could further explain the late oN1 pattern of results, where the activation in the P1T condition is largest because participants experience reliable action–outcome associations on many trials, followed by the P2T condition in which reliable associations are also experienced, but half as much (half of the trials for each tone), while in the U2T condition, little to no reliable activation was observed, because stable action–effect associations could not be established. On the other hand, given the high similarity between the temporal dynamics of the early and late oN1 results that we report here, it is likely that they represent the same process corresponding to the second mechanism of the model only, namely detection of prediction error reflected at different levels along the processing hierarchy, rather than different processes. This is also because the signal “sharpening” that we report here for predicted events is not preparatory as postulated by the initial process of the model, but happens at the time when the stimulus would have been processed if it were there.

Nevertheless, the fact that the simultaneous action-effect predictions are represented later at the level of the late oN1 component (regardless if attention, the opposing process theory, or similar mechanisms explain these results) points to a system that is highly flexible. Likely, it is more efficient like this, given the task/context. It could be that with more practice in which the specific associations for the different key-presses are strengthened and consolidated, the omission-related activity would arise earlier and stronger. How would the brains of more experienced piano players process the omitted tones if the piano were broken, by contrast to the brains of less experienced players?

### 4.3 No costs at later processing levels

Thus far, we have argued that processing costs between predicting effector-specific vs. effector-unspecific, and unique vs. simultaneous tones play a role at early sensory levels. At later processing levels, however, this does not seem to be the case. That is, the data show that early (256 ms) and late (334 ms) oP3 responses are strongly elicited to omissions in all three conditions, while there seem to be no condition differences. Thus, a rather more general sound representation seems to be indexed at this level.

The evidence coming from auditory oddball studies suggests that the P3 component reflects high-level processes driven by an involuntary attention switch, following deviations that exceed a certain threshold which is associated with elicitation of the N1 and MMN components (for a review, see Näätänen, Paavilainen, Rinne, & Alho, 2007). It has also been shown that the P3 elicitation can occur without prior N1-MMN effects (Horváth, Winkler, & Bendixen, 2008), for example following stimuli that are socially or emotionally significant (Bobes, Quínonez, Perez, Leon, & Valdés-Sosa, 2007). In line with this, Nieuwenhuis, De Geus, and Aston-Jones (2011) proposed that the P3 component indexes cognitive processing following motivationally significant stimuli, representing the central nervous system’s counterpart to the orienting response of the sympathetic nervous system. While it is unclear if the omission P3 and the P3 responses following exogenous stimulation reflect similar processes, it seems that here, the presentation of a tone on every trial, rather than which specific tone, is the relevant aspect for the system. In other words, at this processing level, the “when” and “whether” components of intentional action might represent the crucial aspect, rather than the “what” component (Brass & Haggard, 2008; Kriegerhoff, Brass, Prinz, & Waszak, 2009). Alternatively, it is possible that in the effector-unspecific condition, participants learnt binary (by contrast to unique or hand-specific) associations, that is they learnt that both actions generate both tones—in this context, the “what” component plays a role that only becomes evident at the oP3 processing stage. It remains, however, for future research to address the individual and joint contributions of “what,” “when” and “whether” in the auditory processing hierarchy, inclusive of stimulus omissions.
Finally, omission-related designs have been used rather infrequently to investigate action-related predictions; however, the body of omission studies looking at prediction-related activity outside the context of action and for different sensory modalities, is considerably larger. While it is beyond our scope to offer a full review, we briefly point to the results of Bendixen et al., which suggest that the auditory system preactivates the representations of the expected input. Specifically, in sequences of tones consisting of pairs in which the second tone was a repetition of its predecessor, omitting the second (predictable) tone lead to larger omission-related ERPs, by contrast to omitting the first tone for which no specific representation could be formed (Bendixen, Schröger, & Winkler, 2009). Similarly, fMRI results from the visual domain indicate that omissions of expected Gabor patches having certain orientations induce feature-specific activation in the primary visual cortex (similarly to that evoked by the actual stimuli). Importantly, the activation was larger in voxels that prefer the orientation in question, than in other voxels preferring other orientations (Kok, Failing, & de Lange, 2014). This is further supported by MEG data, which indicates that the stimulus-specific sensory templates were already active before the stimulus was presented (Kok, Mostert, & De Lange, 2017).

Our results are congruent with the idea that the activation at early sensory levels following unexpected omissions varies depending on the specificity of the representations. That is, we observe higher activation in the P2T condition where the expected stimulus can be determined precisely, by contrast to the U2T condition, where any of two stimuli can be expected. However, our results additionally point out that in the case of action-related predictions, specificity might be less important at later processing levels—this is reflected in the oP3, which, unlike the earlier components, is similarly elicited across conditions. To conclude, as recently suggested, predictions resulting from action might indeed be special (Dogge, Hofman, Custers, & Aarts, 2019), potentially due to the several components of intentional action operating at different levels (Brass & Haggard, 2008).

### 4.4 Where is the oN2?

Surprisingly, we do not find any oN2 effects, by contrast to previous studies looking at action-related omission responses (Dercksen et al., 2020; SanMiguel, Saupe, et al., 2013; SanMiguel, Widmann, et al., 2013). We find instead a late oN1 component, whose peak (152 ms) closely corresponds to the mean of the time windows/peaks identified for the oN2 analysis in previous studies (116–182 in SanMiguel, Widmann, et al., 2013, and 144–164 ms in SanMiguel, Saupe, et al., 2013, 168 ms in Dercksen et al., 2020). However, based on the topographical maps displaying negative activation strictly around the temporal electrodes (see Figure 3) and not around frontocentral electrodes like it would be typical for an N2 response (for a review, see Folstein & Van Petten, 2008), we can only conclude that the observed component is part of the oN1 response. Two alternative explanations could justify this pattern of results and differences to previous studies.

First, as we discussed earlier, it is possible that due to the rather difficult task requirements, the attentional resources were diverted away from the processing of sounds and omissions, to performing the task itself. It has been proposed that the N2 elicitation reflects a cognitive control mechanism that involves “strategic monitoring and control of motor responses,” where active attention to the stimuli of interest plays a decisive role (Folstein & Van Petten, 2008). Thus, even though it is unclear whether the oN2 can be described in the same way, it seems possible that under circumstances whereby attention allocation to the tones and omissions in the current task design is poor, no oN2 responses are elicited.

Alternatively, the explanation might lie in different pre-processing strategies of the EEG data. Specifically, SanMiguel et al. used rather high high-pass filter cut-off frequencies, that is 0.5 Hz in SanMiguel, Widmann, et al. (2013), and 1 Hz in SanMiguel, Saupe, et al. (2013), respectively. Dercksen et al. (2020) used a much lower high-pass filter cut-off frequency (0.1 Hz); however, they interpret the activity between 100 and 200 ms as omission MMN (oMMN), and not as an oN2. Several methodological papers showed that filtering the data with high thresholds for the high-pass filter leads to significant distortions in the data (Acunzo, MacKenzie, & van Rossum, 2012; Tanner, Morgan-Short, & Luck, 2015; Widmann et al., 2015). Specifically, in the context of an auditory oddball paradigm, Widmann et al. (2015) showed that for the deviant wave, the peak amplitudes of the N1/MMN and N2 components were artificially enhanced, while the P3 component was reduced. Thus, excessive high-pass filtering can induce misleading effects in the N1-N2 time windows by projecting inverted activity from the subsequent P3 component. It is possible that such filtering distortions could also occur for auditory processing of omissions (please refer to the Supporting Information for more details). Finally, note that the oN2/oMMN topographies reported by SanMiguel, Saupe, et al. (2013) and Dercksen et al. (2020), respectively, also show negativity over the temporal areas, in addition to the frontocentral one—this might thus correspond to what we observe and report here as a late oN1, overlapped by an oN2 and/or possibly by filter artefacts.

### 5 Conclusion

The present results extend on previous work emphasizing the fundamental contribution of top-down expectations...
on action-effect predictions. Crucially, we show that the specificity of action-effect associations modulates the sensory processing of omissions. That is, effector-un-specific (by contrast to effector-specific) and simultaneous (by contrast to unique) associations seem to involve some processing costs reflected in the oN1 response. Nevertheless, at later processing levels associated with the oP3 component, it seems that a general sound representation mechanism operates, for which “when” and “whether” might be more important than “what.”

ACKNOWLEDGEMENTS
This work was supported by the German Research Foundation (www.dfg.de) with the grant number SCHR 375/25-1. We would like to thank Taavi Wenk for his help with the data collection, Dr. Benjamin Toovey for proof-reading this manuscript, and the reviewers for helping to improve the quality of this manuscript and providing valuable discussion points.

CONFLICTS OF INTEREST
The authors declare no conflicting interests.

AUTHOR CONTRIBUTIONS
BK, ES and AW designed the study. BK and AW programmed the experimental task. BK was responsible for data collection and analysis. BK wrote the manuscript draft. AW and ES provided feedback and contributed to revisions of the manuscript.

DATA AVAILABILITY STATEMENT
The EEG PCA solution can be accessed on the Open Science Framework at https://osf.io/d245v/. All other data including stimuli, analysis code and raw EEG are available upon request.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/ ejn.14899.

ORCID
Betina Korka https://orcid.org/0000-0002-2702-9108
Erich Schröger https://orcid.org/0000-0002-8321-6629
Andreas Widmann https://orcid.org/0000-0003-3664-8581

REFERENCES
Acunzo, D. J., MacKenzie, G., & van Rossum, M. C. (2012). Systematic biases in early ERP and ERF components as a result of high-pass filtering. Journal of Neuroscience Methods, 209(1), 212–218.
Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. Trends in Cognitive Sciences, 16(7), 390–398.
Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. International Journal of Psychophysiology, 83(2), 120–131.
Bendixen, A., Schröger, E., & Winkler, I. (2009). I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. Journal of Neuroscience, 29(26), 8447–8451.
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.
Bobes, M. A., Quiñonez, I., Perez, J., Leon, I., & Valdés-Sosa, M. (2007). Brain potentials reflect access to visual and emotional memories for faces. Biological Psychology, 75(2), 146–153.
Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. The Neuroscientist, 14(4), 319–325.
Chatrain, 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(2), 83–92.
Chaumon, M., Bishop, D. V. M., & 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.
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.
Derksen, T. T., Widmann, A., Schröger, E., & Wetzel, N. (2020). Omission related brain responses reflect specific and unspecific action-effect couplings. NeuroImage, 215, 116840.
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(1), 138–145.
Dogge, M., Hofman, D., Custers, R., & Aarts, H. (2019). Exploring the role of motor and non-motor predictive mechanisms in sensory attenuation: Perceptual and neurophysiological findings. Neuropsychologia, 124, 216–225.
Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H. J., Kirsch, H. E., Knight, R. T., & Deouell, L. Y. (2019). Direct evidence for prediction signals in frontal cortex independent of prediction error. Cerebral Cortex, 29(11), 4530–4538.
Eaton, J. W., Bateman, D., Hauberg, S., & Wehbring, R. (2016). GNU Octave: A High-level Interactive Language for Numerical Computations: Edition 4 for Octave version 4.0.0. Boston, MA: Free Software Foundation.
Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. Journal of Experimental Psychology: Human Perception and Performance, 27(1), 229.
Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. Frontiers in Human Neuroscience, 4, 215.
Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. Psychophysiology, 45(1), 152–170.
Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of The Royal Society of London B: Biological Sciences, 360(1456), 815–836.
Friston, K. (2018). Does predictive coding have a future? Nature Neuroscience, 21(8), 1019–1021.
Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. Science, 182(4108), 177–180.
Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. Brain Research, 1626, 54–65.

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

Hughes, G., Desantis, A., & Waszak, F. (2013a). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. Psychological Bulletin, 139(1), 133–151.

Hughes, G., Desantis, A., & Waszak, F. (2013b). Attenuation of auditory N1 results from identity-specific action-effect prediction. European Journal of Neuroscience, 37(7).

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. Perception, 36(14), 1.

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. Journal of Cognitive Neuroscience, 26(7), 1546–1554.

Lee, M. D., & Wagemakers, E.-J. (2013). Bayesian cognitive modeling: A practical course. Cambridge, UK: Cambridge University Press.

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. Journal of the Optical Society of America A, 20(7), 1434–1448.

Martikainen, M. H., Kaneko, K. I., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. Cerebral Cortex, 15(3), 299–302.

Marzecóva, A., Schettino, A., Widmann, A., SanMiguel, I., Kotz, S. A., & Schröger, E. (2018). Attentional gain is modulated by probabilistic feature expectations in a spatial cuing task: ERP evidence. Scientific Reports, 8(1), 1–14.

Mill, R., Böhm, T., Bendixen, A., Winkler, I., & Denham, S. L. (2011, March). 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). Baltimore, MD: IEEE.

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

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(2), 162–175.

Pieszek, M., Widmann, A., Gruber, T., & Schröger, E. (2013). The human brain maintains contradictory and redundant auditory sensory predictions. PLoS One, 8(1), e53634.

Press, C., Kok, P., & Yon, D. (2020). The perceptual prediction paradox. Trends in Cognitive Sciences, 24(1), 13–24.

Press, C., & Yon, D. (2019). Perceptual prediction: Rapidly making sense of a noisy world. Current Biology, 29(15), R751–R753.

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., 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(20), 8633–8639.

Schröger, E., Marzecóva, A., & Sanmiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. European Journal of Neuroscience, 41(5), 641–664.

Stuckenberg, M. V., Schröger, E., & Widmann, A. (2019). Presentation probability of visual—Auditory pairs modulates visually induced auditory predictions. Journal of Cognitive Neuroscience, 31(8), 1110–1125.

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. Psychophysiology, 52(8), 997–1009.

Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. BMC Neuroscience, 14(1), 2.

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data – A practical approach. Journal of Neuroscience Methods, 250, 34–46.
Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, 9(1), 1–8.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Korka B, Schröger E, Widmann A. What exactly is missing here? The sensory processing of unpredictable omissions is modulated by the specificity of expected action-effects. *Eur J Neurosci.* 2020;52:4667–4683. [https://doi.org/10.1111/ejn.14899](https://doi.org/10.1111/ejn.14899)