



# Similarities and Differences between Performers and Observers in Processing Auditory Action Consequences: Evidence from Simultaneous EEG Acquisition

Marta Ghio<sup>1</sup>, Sophie Egan<sup>2</sup>, and Christian Bellebaum<sup>2</sup>

## Abstract

■ In our social environment, we easily distinguish stimuli caused by our own actions (e.g., water splashing when I fill my glass) from stimuli that have an external source (e.g., water splashing in a fountain). Accumulating evidence suggests that processing the auditory consequences of self-performed actions elicits N1 and P2 ERPs of reduced amplitude compared to physically identical but externally generated sounds, with such reductions being ascribed to neural predictive mechanisms. It is unexplored, however, whether the sensory processing of action outcomes is similarly modulated by action observation (e.g., water splashing when I observe you filling my glass). We tested 40 healthy participants by applying a methodological approach for the simultaneous EEG recording of two persons: An observer observed button presses executed by a

performer in real time. For the performers, we replicated previous findings of a reduced N1 amplitude for self- versus externally generated sounds. This pattern differed significantly from the one in observers, whose N1 for sounds generated by observed button presses was not attenuated. In turn, the P2 amplitude was reduced for processing action- versus externally generated sounds for both performers and observers. These findings show that both action performance and observation affect the processing of action-generated sounds. There are, however, important differences between the two in the timing of the effects, probably related to differences in the predictability of the actions and thus also the associated stimuli. We discuss how these differences might contribute to recognizing the stimulus as caused by self versus others. ■

## INTRODUCTION

Sensory outcomes of voluntary actions (e.g., self-generated sounds) are perceived as less intense than externally generated stimuli that are otherwise physically identical (Reznik & Mukamel, 2019; Horváth, 2015; Schröger, Marzecová, & SanMiguel, 2015; Crapse & Sommer, 2008). According to sensorimotor control accounts explaining this sensory attenuation, an efference copy of the motor command is fed into a forward model, which computes the expected sensory stimulation. The attenuation then results from the integration of the expected and the actual sensory input (Wolpert & Flanagan, 2001; Wolpert & Ghahramani, 2000). Evidence from neuroimaging and (virtual) lesion studies indeed shows that the sensory modulation by action depends on cortical motor regions that subserve action planning and are a likely source of efference copies (Timm, SanMiguel, Keil, Schröger, & Schönwiesner, 2014; Waszak, Cardoso-Leite, & Hughes, 2012), as well as on the cerebellum, which is considered to provide the computations for the forward model (for reviews see Dogge, Custers, & Aarts, 2019; Reznik & Mukamel, 2019; Crapse & Sommer, 2008).

Current theories of action representation and social cognition suggest that we also predict the sensory consequences of others' actions (Brown & Brüne, 2012). Based

on the finding that mirror neurons fire both for action execution and for the observation of the same action by others, further research has confirmed that the neural mechanisms for action execution and observation overlap (Rizzolatti & Sinigaglia, 2010). Consequently, it has been hypothesized that predictive accounts of sensorimotor processing can be generalized to action observation (Kilner, Friston, & Frith, 2007; Miall, 2003; Wolpert, Doya, & Kawato, 2003). Within the framework of sensorimotor control accounts, it has been claimed that observers may covertly simulate the motor commands for the observed action (which generates a signal similar to efference copies; Wolpert et al. 2003). These are then fed into forward models that generate predictions about observed action outcomes (Wolpert & Flanagan, 2001).

Increasing evidence for the brain's ability to predict observed action consequences comes from psychophysical studies showing prospective eye movements when observing grasping movements (Rotman, Troje, Johansson, & Flanagan, 2006; Flanagan & Johansson, 2003) and from observational motor learning research (Wolpert, Diedrichsen, & Flanagan, 2011). Studies on motor experts (e.g., athletes, dancers) indicate that the capacity to predict action outcomes can depend on our motor skills (Aglioti, Cesari, Romani, & Urgesi, 2008; Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006), in line with the simulation hypothesis that others' actions can be decoded by activating

<sup>1</sup>University of Trento, <sup>2</sup>Heinrich Heine University Dusseldorf

one's own action system at a subthreshold level (Keysers & Gazzola, 2007; Decety & Grèzes, 2006). A controversial issue, however, is whether the perceptual processing of sensory action outcomes is modulated by action observation in a similar way as by action execution. Conflicting evidence comes from perceptual research on a potential sensory attenuation effect for auditory stimuli following observed actions. While Sato (2008) described a sensory attenuation effect for action observation, this finding could not be confirmed in later studies (Weiss & Schütz-Bosbach, 2012; Weiss, Herwig, & Schütz-Bosbach, 2011a, 2011b). Cao and Gross (2015) also reported mixed results, which were explained in terms of cultural differences between eastern (more interdependent) and western (more independent) cultures based on a negative correlation between sensory attenuation by observation and independent self-construal scores.

To gain insights into the neural mechanisms involved in the processing of auditory outcomes of actions, EEG has been applied. Studies on active performance showed that the amplitude of both the N1 and P2 ERP components (or their magnetic equivalents) is reduced for self- versus externally generated sounds (Reznik & Mukamel, 2019; Horváth, 2015). These components appear to be functionally dissociated. The N1 reduction, but not the P2 reduction, is absent in cerebellar lesion patients, suggesting that it reflects cerebellar forward model predictions based on motor information (Knolle, Schröger, & Kotz, 2013). The role of the forward model and efference copies in predicting environment-related action outcomes such as tones has recently been questioned (Dogge et al., 2019). Instead, the predictive coding account, which also claims a modulatory role of action outcome anticipation in perception, has suggested different prediction mechanisms (Picard & Friston, 2014; Pickering & Clark, 2014). According to this account, motor-related information is only one potential source for predictions, and the prediction effects are less specific than those ascribed to forward models. However, for tones elicited at or near action onset as in the study by Knolle et al. (2013), there is reason to believe that predictions derived from motor-related information available before the action, and thus probably efference copies, play an important role, because, otherwise, the predictions could not affect the early processing of action outcomes in the N1 time window. As Dogge et al. (2019) have pointed out, forward model predictions may also contribute to the processing of environment-related outcomes, especially when the timing of the action outcome is important. The P2 reduction, in turn, seems to be independent from the cerebellum (Cao, Veniero, Thut, & Gross, 2017; Knolle et al., 2013) and may thus reflect higher order and more general prediction mechanisms in the sense of predictive coding accounts.

The timing of prediction generation is thus a critical issue in the processing of outcomes of voluntary actions and should be carefully considered for action observation. Focusing on when motor regions are involved in action

observation, there is an important difference compared to active performance. When the onset time of the observed action cannot be predicted with high precision, motor areas in the observer are activated only after the action onset (Sebastiani et al., 2014; Nishitani & Hari, 2000). In such a case, motor-related information in the form of an efference copy is not available before the movement, and thus, also the prediction generation cannot take place. In a previous EEG study, we thus hypothesized that efference copies play no or only a minor role in the processing of sensory outcomes of observed actions, which would result in an absent or at least a weaker N1 amplitude reduction compared to action performance (Ghio, Scharmach, & Bellebaum, 2018). For later processing as reflected in the P2, altered processing such as an amplitude reduction was considered conceivable in action observation, as reflecting more general prediction mechanisms. However, we did not find a clear dissociation between the N1 and the P2, with indication of N1 amplitude reduction also in observers (Ghio et al., 2018). In that study, however, the timing of performed and observed action was not perfectly matched. We used an animation on a computer screen for action observation, which entailed a delay between the first indication of the action and its auditory outcome, and this might have led to the observed N1 amplitude reduction in observers.

In this study, we therefore applied a methodological approach for the simultaneous EEG recording in two participants (Babiloni & Astolfi, 2014): One participant directly observed button presses performed by another participant in real time, while brain activity was measured in both to assess the processing of sounds as action outcomes. Because of the low predictability of the exact timing of the observed action, we expected that action-related information would not be available before the observed movement so that the N1 for action-generated sounds would be reduced only for performed actions, resulting in larger amplitudes for observers than performers. Concerning later processing in the P2 time window, we expected an amplitude reduction for action-generated sounds relative to external sounds for both action performance and observation.

## METHODS

### Participants

Forty participants (15 men; mean age = 23.3 years,  $SD = 5.0$  years), who were equally distributed on a performer and an observer condition (see below), took part in the experiment. An a priori power analysis was conducted with G\*Power 3.0 (Faul, Erdfelder, Lang, & Buchner, 2007) to determine the required sample size for the within-between interaction effect that we planned to test by applying a mixed ANOVA (see below). Given an alpha value of .05 and an effect size of .25 (that corresponds to an effect size of .61 described by Cohen, 1988), 36 participants were required to achieve a power of .90.

Participants were recruited two by two (total of 20 pairs). For each pair, we randomly assigned one participant to the role of the performer and the other one to the role of the observer, yielding one group of 20 performers (eight men; mean age = 23.2 years,  $SD = 4.1$ ; 18 right-handed, one left-handed, one bimanual) and one group of 20 observers (seven men; mean age = 23.5,  $SD = 5.8$ ; 17 right-handed, three left-handed). The two groups differed neither in age,  $t(38) = -0.157, p = .876, d = -0.050$ ; nor sex, Pearson  $\chi^2(1) = 0.107, p = .744$ , Cramer's  $V = 0.052$ ; nor handedness, Pearson  $\chi^2(2) = 2.029, p = .363$ , Cramer's  $V = 0.225$ . All participants reported normal hearing, normal or corrected-to-normal visual acuity, and no history of neurological or psychiatric disorders or medication use. All participants gave written informed consent and received either €15 or course credit for their participation. The ethics committee of the Faculty of Mathematics and Natural Sciences at Heinrich Heine University Düsseldorf, Germany, has approved this study.

## Experimental Design

We adapted the block-designed self-generation paradigm (for a review see Horváth, 2015) to test the processing of the auditory sensory consequences not only of self-performed actions but also of observed actions, with the EEG data from the performer and the observer being simultaneously recorded during the entire duration of the experiment (for a schematic overview, see Figure 1). Specifically, the processing of action-generated sounds (ACT) elicited by either self-performed or observed button presses was compared to the processing of physically identical, externally generated sounds. By applying the variant of the self-generation

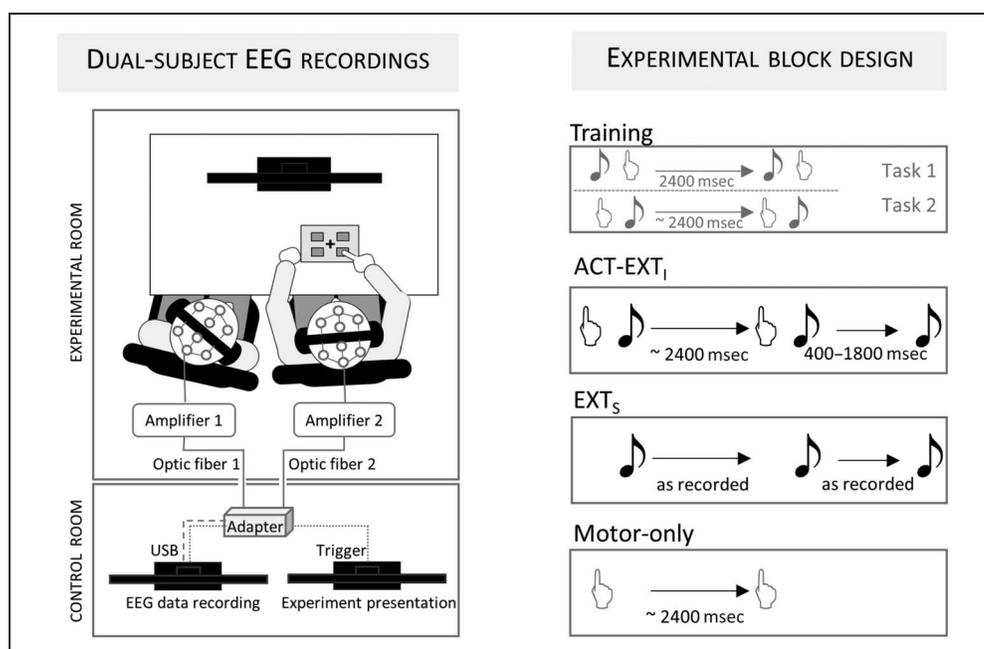
paradigm from Knolle et al. (2013) and Baess, Horváth, Jacobsen, and Schröger (2011), externally generated sounds were presented either intermixed within the action-outcome block (EXT<sub>I</sub>) or in a separate block (EXT<sub>S</sub>). This variant of the paradigm has two advantages. First, it allows a within-block comparison between ACT and EXT<sub>I</sub>, thus avoiding potential biases because of differences between blocks (Baess et al., 2011). Second, it allows the comparison with previous studies applying the standard self-generation paradigm, which comprises the ACT and EXT<sub>S</sub> only. To sum up, our mixed factorial design included the within-subject factor Sound (ACT, EXT<sub>I</sub>, EXT<sub>S</sub>) and the between-subject factor Group (performers, observers). As in the standard self-generation paradigm, this variant also comprised an initial training block—to ensure that the participants learned the specific (observed) action-auditory outcome association—and a motor-only block—to control for effects of motor activity associated with the (observed) button press on the ERPs.

## Experimental Blocks

### Training Block

The aim was to train performers to press a button in a regular rhythm (one every  $2400 \pm 600$  msec; see Ghio et al., 2018; Knolle et al., 2013) and to familiarize them and the observers with the sound that accompanied each button press. The training comprised two tasks. In the first task, performers tried to press the button simultaneously to sounds that were presented every 2400 msec. In the second task, performers produced a sequence of sound-generating button presses in the previously learned rhythm. Button presses were considered as correct, when

**Figure 1.** *Left:* Setting for the simultaneous EEG recording. Both participants wore an electrode cap; the signal was measured by using separate amplifiers with separate ground electrodes for each individual. Each amplifier was optically coupled to an adapter, which then sent all the signals to a single workstation devoted to data acquisition. *Right:* a schema of the variant of the self-generation paradigm applied in the current study. This block-designed paradigm comprises a training block, one block of action-generated sounds elicited by button presses executed by the performer and intermixed with externally generated sounds (ACT-EXT<sub>I</sub>), one block including externally generated sounds only (EXT<sub>S</sub>), and a motor-only block.



they occurred in an interval between 1800 and 3000 msec after the previous button press. If participants performed the button presses outside of this time window, they received written feedback informing them whether their response was too early or too late (Knolle et al., 2013). Both tasks consisted of 150 trials (two separate subblocks, each of 75 consecutive button presses with the right and the left index finger, respectively). During the training, observers were instructed to merely observe the performer's button presses and to listen to the sounds. To reduce eye movements that affect the quality of the registered EEG signal, active performers were asked to look at the fixation cross on the computer screen. Observers were asked to look at a fixation cross that was placed on the response box, in a position that enabled the participants to observe the button presses of the performers without eye movements. To continue with the experiment, the performer had to correctly perform in 75% of the trials in the second task. All pairs fulfilled the criterion at first try.

#### *ACT-EXT<sub>1</sub> Block*

Performers pressed a button to generate ACT sounds in the same rhythm they learned in the training block and were instructed to carefully listen to them. In addition, EXT<sub>1</sub> were randomly delivered by the computer in 40% of the trials. EXT<sub>1</sub> followed the ACT with a delay of either 400, 800, 1200, or 1800 msec (each delay was presented in 10% of the trials). One ACT-EXT<sub>1</sub> block consisted of two subblocks of 75 button presses performed with the left and right index finger, respectively, yielding 150 ACT and 60 EXT<sub>1</sub> trials. The ACT-EXT<sub>1</sub> block was repeated, for a total of 300 ACT and 120 EXT<sub>1</sub> trials. The observers were asked to observe the button presses and to listen to the sounds without performing any action themselves.

#### *EXT<sub>S</sub> Block*

Following the procedure described in previous work (Ghio et al., 2018; Knolle et al., 2013; Baess et al., 2011), both ACT and EXT<sub>1</sub> of the preceding ACT-EXT<sub>1</sub> block were played back in an EXT<sub>S</sub> block. Only the playbacks of ACT, however, were considered as EXT<sub>S</sub> in the following analyses to ensure that ACT and EXT<sub>S</sub> were similarly predictable (Ghio et al., 2018; Knolle et al., 2013). Accordingly, there were two EXT<sub>S</sub> blocks, each including 150 EXT<sub>S</sub> trials (300 in total). The instructions for both the performers and the observers were to carefully listen to the sounds without performing any action.

#### *Motor-only Block*

Performers pressed the button in the learned 2400-msec interval (two subblocks, each of 75 consecutive button presses with the right and the left index finger, respectively, 150 in total). Importantly, button presses did not generate sounds. Observers were instructed to observe the button presses

without performing any action themselves. Because an EXT<sub>S</sub> block always had to follow an ACT-EXT<sub>1</sub> block, the motor-only block was presented either before, between, or after the two iterations of the other blocks, to counterbalance the block presentation order across participants.

## **Data Acquisition**

### *Setup*

The performer and the observer were comfortably seated next to each other in an electrically and acoustically shielded experimental room in front of a BenQ (EW2740L, 27 inch LED, full HD) computer screen with a screen resolution of 1920 × 1080 and a 60-Hz refresh rate (Figure 1). The response box (Response Pad Model RB-844 from Cedrus Corporation; www.cedrus.com) was placed between the screen and the performer. Presentation software (Version 17.2, Build 10.08.14) from Neurobehavioral Systems Inc. (www.neurobs.com) was used for the presentation of all stimuli. Visual stimuli, namely, instructions and fixation cross, were presented in white letters on a black background. The 680-Hz sound lasted 50 msec, with an increase of amplitude in the first and a decrease in the last 5 msec. This sound was created with MATLAB (R2014a, The MathWorks, Inc.) and delivered binaurally via headphones (Sennheiser HD), with the same intensity for all participants. By performing measurements with a Tectronix TDS 210 Two Channel Digital Real Time Oscilloscope (60 MHz, 1 GSample/sec), we verified that ACT onset was 50 msec after the button press was registered by the Cedrus box, and in the time interval between the button press and the button release (Ghio et al., 2018).

### *Simultaneous EEG Acquisition*

EEG data were simultaneously recorded from the pair (Babiloni & Astolfi, 2014). Both the performer and the observer wore an Easycap electrode cap (www.easycap.de), on which we attached passive Ag/AgCl electrodes according to the international 10/20 system. Electrodes were positioned at the following scalp sites: F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, PO7, PO3, POz, PO4, PO8. For each participant, the electrodes on the mastoids (TP9, TP10) served as reference electrodes, and the ground electrode was attached to the AFz position. Four electrodes were used to monitor vertical (position FP2 and one electrode under the right eye) and horizontal (positions F9 and F10) eye movements. To simultaneously record the signal from both participants, separate Brainamp amplifiers with separate ground electrodes were used for each individual. Each amplifier was optically coupled to a BrainVision USB2 adapter, which then sent all the signals to a single workstation devoted to data acquisition. On the workstation, data of each participant were recorded by using BrainVision Recorder (Version 1.20.0506, Brain Products GmbH; www.brainproducts.com).

The signal was recorded at a sampling rate of 1000 Hz. Impedances were kept below 5 k $\Omega$ , and direct-current corrections were applied on-line, when necessary.

## Data Analysis

Statistical analysis was conducted with JASP (JASP Team, 2019). For all inferential statistics, an alpha level of .05 was declared. Degrees of freedom were adjusted according to Greenhouse–Geisser, when violations of sphericity occurred. Planned follow-up tests for significant interactions as well as multiple correlations were corrected for the false discovery rate (FDR), with the procedure introduced by Benjamini and Hochberg (1995). As measure of effect size, we report  $\eta_p^2$  or Cohen's  $d$ , where appropriate.

### ERP Components

**Preprocessing.** The same preprocessing procedure was applied for both performers' and observers' data using the BrainVision Analyzer software (Version 2.1.0.327, Brain Products) and MATLAB (2014Ra). We applied a global direct current detrend, Butterworth zero phase filters (low cutoff = 0.3 Hz, 12 dB/oct; high cutoff = 15 Hz, 12 dB/oct) and a 50-Hz Notch-Filter (see Baess et al., 2011, and Knolle et al., 2013, for a similar filtering procedure). Artifacts because of blinks and horizontal eye movements were detected and discarded for each participant by using the ocular correction independent component analysis (ICA; ICA steps = 512, Infomax restricted biased) as implemented in BrainVision Analyzer.

ERPs were time-locked to the onset of ACT, EXT<sub>I</sub>, and EXT<sub>S</sub>. For ACT, we included sounds generated by both left and right button presses. The sounds generated by button presses in an interval shorter than 1800 msec or longer than 10000 msec, or by a double button press were excluded from further analyses. The EXT<sub>I</sub> of the different delays after ACT were pooled for the analysis (as in Ghio et al., 2018; Knolle et al., 2013). Each epoch had a duration of 600 msec, including a 150-msec presound period, and was baseline-corrected using the mean amplitude in the time window between 150 msec and 50 msec before the sound. Epochs were rejected if they contained artifacts, which were detected by applying an automatic algorithm as implemented in BrainVision Analyzer, with the following customized parameters: maximal allowed voltage step = 50  $\mu$ V/msec, maximal allowed difference of values within 100-msec intervals = 100  $\mu$ V, maximal/minimal allowed amplitude =  $\pm$ 100  $\mu$ V, lowest activity of 0.5  $\mu$ V within 100-msec intervals. Then, the average signal was created for ACT, EXT<sub>I</sub>, and EXT<sub>S</sub>, respectively.

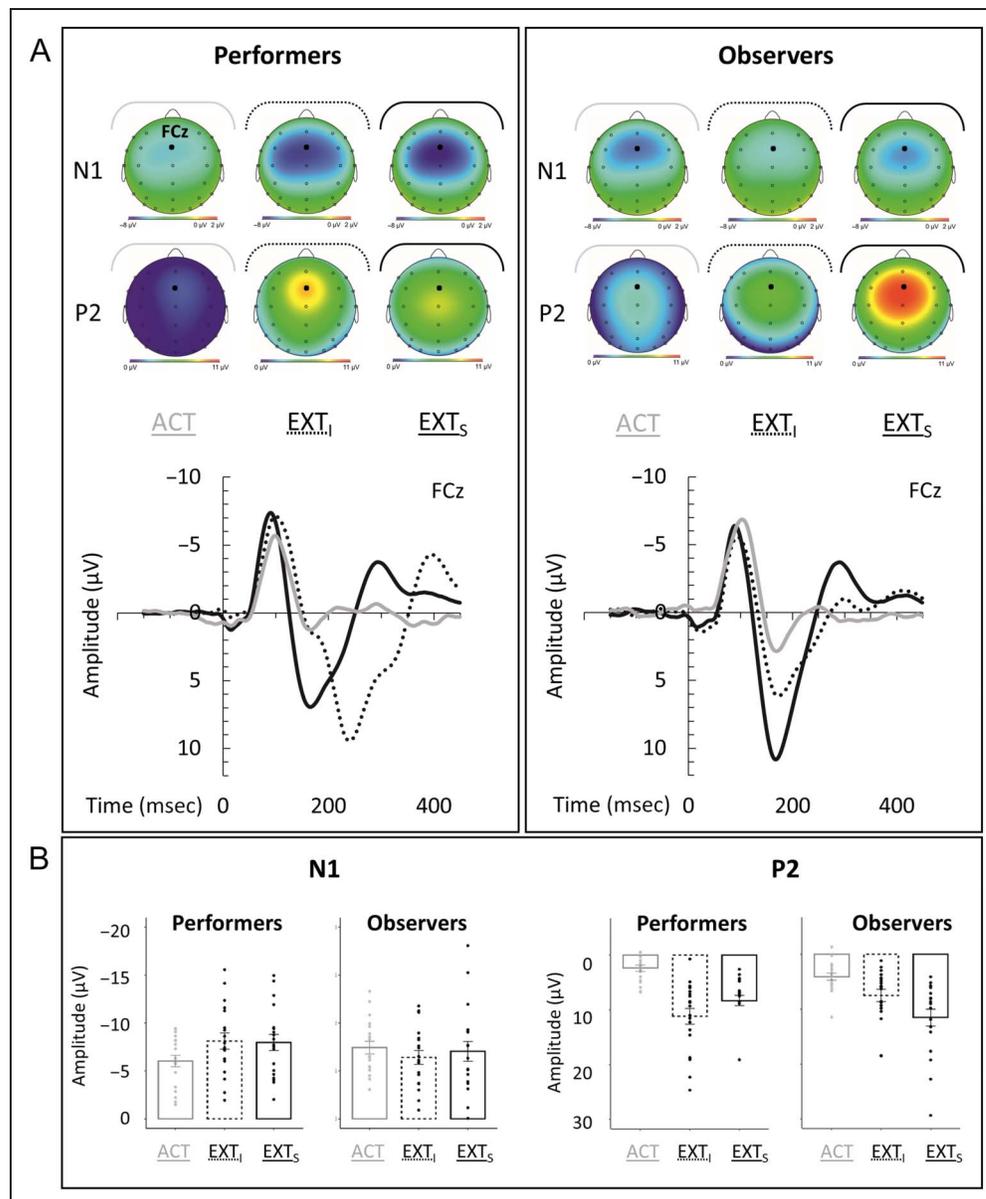
To control for the motor activity associated with generating ACT, we applied the standard procedure for the self-generation paradigm, which consists in subtracting the ERPs of the motor-only block from the ERPs associated with ACT (Horváth, 2015). Before applying such a correction, we verified that the mean intervals between button presses generating ACT ( $M = 3176$  msec,  $SD = 1130$

and the button presses in the motor-only condition ( $M = 3071$  msec,  $SD = 960$ ) did not differ significantly,  $t(18) = 0.519$ ,  $p = .610$ ,  $d = 0.119$ . This is an important prerequisite for applying the motor correction to the auditory ERPs, because differences in action timing between button presses for inducing the sounds and button presses in the motor-only condition could be indicative of action-related processing differences (Horváth, 2015). Then, we created ERPs time-locked to the (observed) button press in the motor-only condition, with epoch duration of 600 msec ( $-100$  to 500 msec relative to the marker of the button press). After rejecting trials with artifacts (see above), we calculated the average ERPs for the motor-only block and subtracted it from the ERPs of the ACT. In all following analyses, the ACT ERPs refer to these corrected values.

**Statistical analyses.** Participants whose data set contained less than 70% of the epochs for either ACT, EXT<sub>I</sub>, or EXT<sub>S</sub> or the motor-only ERPs, because of either reduced accuracy (concerning button press interval) of the performer or numerous artifacts, were excluded from analysis. This was the case for one performer and the associated observer. Therefore, the following analyses were conducted on the data of 38 participants: 19 performers (seven men; mean age = 22.8 years,  $SD = 3.9$ ; 17 right-handed, one left-handed, one bimanual) and 19 observers (seven men; mean age = 23.6,  $SD = 5.9$ ; 16 right-handed, three left-handed). The percentage of excluded epochs for ACT (performers:  $M = 6.4\%$ ;  $SD = 7.7\%$ ; observers:  $M = 7.5\%$ ;  $SD = 8.5\%$ ), EXT<sub>I</sub> (performers:  $M = 2.9\%$ ;  $SD = 2.6\%$ ; observers:  $M = 3.4\%$ ;  $SD = 5.7\%$ ), and EXT<sub>S</sub> (performers:  $M = 7.4\%$ ;  $SD = 7.4\%$ ; observers:  $M = 9.2\%$ ;  $SD = 9.1\%$ ) did not differ across groups, as indicated by a not significant main effect of Group,  $F(1,36) = 0.306$ ,  $p = .584$ ,  $\eta_p^2 = .008$ , and Sound  $\times$  Group interaction,  $F(1,169, 42.077) = 0.274$ ,  $p = .640$ ,  $\eta_p^2 = .008$ .

Statistical analyses were conducted on the amplitudes of the N1 and P2 (potential interpretations of this component as a P3a are addressed in the Discussion) ERP components. Based on the visual inspection of the topographies across groups and conditions (see Figure 2A), we observed pronounced N1 and P2 amplitudes and attenuation effects at frontocentral midline electrodes, in accordance with previous studies (e.g., Knolle et al., 2013; Baess et al., 2011; Lange, 2011). Because previous studies did not justify hypotheses concerning differences between frontocentral midline electrodes, we restricted the analyses to the electrode site FCz, which belonged to the electrode cluster with most pronounced amplitudes across all sound conditions and groups for both components. The time windows for the detection of the N1 and the P2 on the single participant level were determined based on visual inspection of the grand average ERPs. Separately for ACT, EXT<sub>I</sub>, and EXT<sub>S</sub>, the N1 and P2 were scored as the maximum negative peak amplitude between 50 and 150 msec and the maximum positive peak amplitude between 100 and 300 msec after sound onset, respectively. For each of the two ERP components, we carried out a mixed 3  $\times$  2 ANOVA, including

**Figure 2.** (A) Topographies for each component (N1 and P2) and grand-averaged ERPs at electrode FCz are displayed separately for performers (on the left) and observers (on the right). Topographies and ERPs are shown for the three levels of the factor Sound: action-generated sounds (ACT, gray line), externally generated sounds in the intermixed condition (EXT<sub>I</sub>, dotted line), and externally generated sounds in the separate condition (EXT<sub>S</sub>, black line). (B) Mean peak amplitudes for N1 (right) and P2 (left) at FCz, separately for the levels of Sound and Group. Filled circles represent single-subject data. Error bars represent  $\pm 1$  SEM.



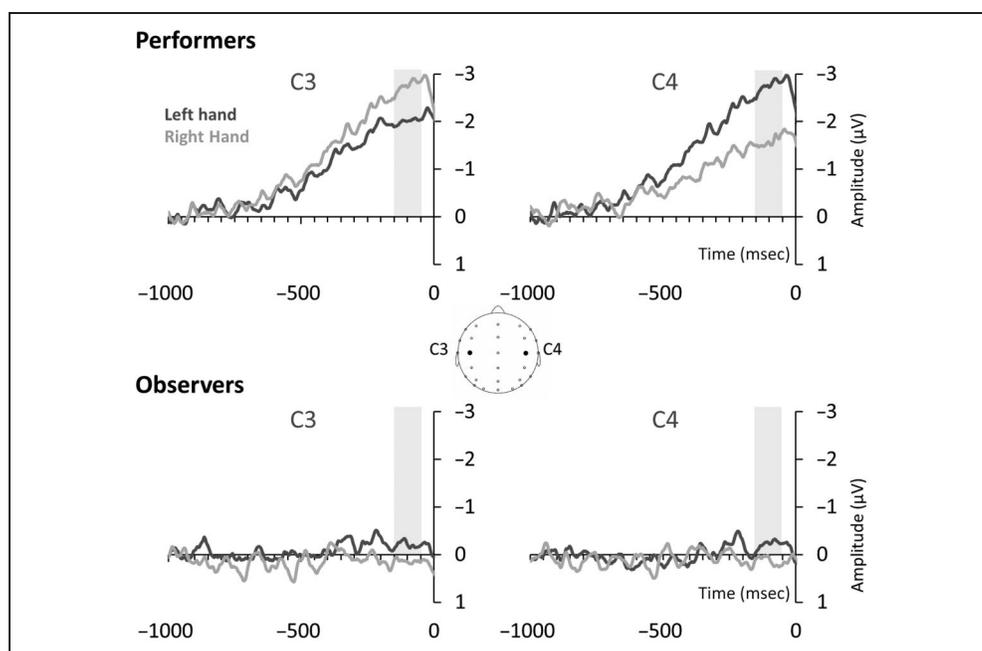
Sound (ACT, EXT<sub>I</sub>, and EXT<sub>S</sub>) as a within-subject factor and Group (performers, observers) as a between-subject factor. Furthermore, visual inspection also revealed differences between performers and observers with respect to the latency in the P2 component associated with the processing of EXT<sub>I</sub>. We thus additionally applied a  $3 \times 2$  (Sound  $\times$  Group) ANOVA on the P2 latencies.

### Preaction Potential

Based on previous studies that linked prediction mechanisms to action preparation processes (e.g., Poonian, McFadyen, Ogden, & Cunnington, 2015; Ford, Palzes, Roach, & Mathalon, 2014), and similar to our previous study (Ghio et al., 2018), we also analyzed EEG signals indicative of action preparation both in performers and observers (preaction or readiness potential; Poonian et al., 2015; Ford et al.,

2014). We applied the same filtering procedure described above, except for the low cutoff for the Butterworth zero phase filters. Specifically, based on the previous work by Poonian et al. (2015; see also Ghio et al., 2018), we used a low cutoff of 0.05 Hz to retain slow-wave activity associated with preaction potentials, while attenuating slower baseline drift. ICA ocular correction was applied as described above. Then, we created epochs time-locked to the marker of the (observed) button presses (starting 1000 msec before and lasting until 500 msec after button press onset, duration 1500 msec) in the ACT-EXT<sub>I</sub> block, separately for right and left index finger button presses. Each epoch was baseline-corrected using the mean amplitude in the time interval between  $-1000$  and  $-900$  msec as visual inspection of the ERP data at C3 and C4 for both performers and observers (see Figure 3) did not show a negative potential during this time interval (for similar epoch definition

**Figure 3.** Preaction potential related to left- and right-hand button press at C3 and C4 electrodes for the performers (top) and the observers (bottom). Gray areas indicate the time windows for amplitude measurements (from  $-150$  to  $-50$  msec).



and baseline correction, see Ford et al., 2014). Button presses in an interval shorter than 1800 msec or longer than 10000 msec, and double button presses were excluded from further analyses. Epochs were rejected if an  $EXT_1$  was included in the time interval between  $-1000$  and  $0$  msec, and if they contained artifacts (see above for criteria). Finally, based on visual inspection, the preaction potentials were defined as the average amplitude in the time interval between  $-150$  and  $-50$  msec before the (observed) button presses performed with either the right or the left hand. Statistical analysis was restricted to electrode sites C3 in the left hemisphere and C4 in the right hemisphere. The mean preaction potentials were entered into a mixed  $2 \times 2 \times 2$  ANOVA, including Hand (left, right) and Hemisphere (left, right) as within-subject factors, and Group (performers, observers) as a between-subject factor.

### Correlation Analyses

The correlation (Pearson's  $r$ ) between N1 reduction and preaction potentials was also examined (see also Ghio et al., 2018; Ford et al., 2014). For this purpose, N1 amplitudes of external sounds ( $EXT_1$  and  $EXT_S$ ) were subtracted from amplitudes of ACT, resulting in two N1 reduction values at the electrode site FCz for each participant. The mean preaction potential was computed as the overall mean across the C3 and C4 electrode and for both the right and the left hand. In addition, for the group of observers, correlations (Pearson's  $r$ ) between the temporal variability of the observed actions, measured as the standard deviation of the timing of ACT responses (Ghio et al., 2018; Weiss et al., 2011a), and N1 amplitudes, as well as N1 reductions, were examined at the electrode site FCz.

## RESULTS

### ERPs

The ERPs for the three types of sounds at the electrode site FCz are displayed in Figure 2A, separately for the performers and the observers. In each group, the N1 and the P2 components are clearly visible in all conditions.

### N1 Component

A  $3 \times 2$  (Sound  $\times$  Group) ANOVA revealed that neither the main effect of Sound,  $F(2, 72) = 1.321, p = .273, \eta_p^2 = .035$ , nor of Group,  $F(1, 36) = 0.170, p = .682, \eta_p^2 = .005$ , was significant. In turn, we observed a significant Sound  $\times$  Group interaction,  $F(2, 72) = 5.539, p = .006, \eta_p^2 = .133$  (see Figure 2B for descriptive statistics). To explain the interaction, we performed a one-way ANOVA with Sound as a within-subject factor separately for each group. For the performers, a significant main effect of Sound was found,  $F(2, 36) = 5.125, p = .011, \eta_p^2 = .222$ . Pairwise comparisons revealed a significant N1 amplitude reduction for ACT compared to both  $EXT_1$ ,  $t(18) = 2.687, p_{FDR} = .041, d = 0.616$ , and  $EXT_S$ ,  $t(18) = 2.407, p_{FDR} = .041, d = 0.552$ . The N1 amplitudes for  $EXT_1$  and  $EXT_S$  did not significantly differ,  $t(18) = -0.274, p_{FDR} = .787, d = -0.063$ . For the observers, the main effect of Sound was not significant,  $F(2, 36) = 1.290, p = .288, \eta_p^2 = .067$ .

### P2 Component

A  $3 \times 2$  (Sound  $\times$  Group) ANOVA revealed a significant main effect of Sound,  $F(2, 72) = 46.917, p < .001, \eta_p^2 = .566$ , with the P2 amplitude for ACT ( $M = 3.228, 95\% CI [1.672, 4.784]$ ) being significantly reduced compared to

EXT<sub>I</sub> ( $M = 9.365$ , 95% CI [7.809, 10.921]) and EXT<sub>S</sub> ( $M = 9.926$ , 95% CI [8.370, 11.482]; both  $p_{FDR} < .001$ ). No difference was found between the P2 amplitudes of EXT<sub>I</sub> and EXT<sub>S</sub> ( $p_{FDR} = .567$ ). The main effect of Group was not significant,  $F(1, 36) = 0.108$ ,  $p = .744$ ,  $\eta_p^2 = .003$ , but we found a significant Sound  $\times$  Group interaction,  $F(2, 72) = 11.191$ ,  $p < .001$ ,  $\eta_p^2 = .237$  (see Figure 2B for descriptive statistics). This interaction was further explored by applying a one-way ANOVA with Sound as a within-subject factor, separately for each group. For the performers, a significant main effect of Sound was found,  $F(1.389, 24.997) = 28.735$ ,  $p < .001$ ,  $\eta_p^2 = .615$ , with the P2 amplitude for ACT being reduced compared to both EXT<sub>I</sub>,  $t(18) = -6.440$ ,  $p_{FDR} < .001$ ,  $d = -1.477$ , and EXT<sub>S</sub>,  $t(18) = -8.620$ ,  $p_{FDR} < .001$ ,  $d = -1.977$ . Moreover, the EXT<sub>I</sub> amplitude was significantly larger than the amplitude for EXT<sub>S</sub>,  $t(18) = 2.110$ ,  $p_{FDR} = .049$ ,  $d = 0.484$ . For the observers, results also revealed a main effect of Sound,  $F(2, 36) = 29.532$ ,  $p < .001$ ,  $\eta_p^2 = .621$ , but with a different underlying pattern. The P2 amplitude for ACT was again reduced compared to both EXT<sub>I</sub>,  $t(18) = -3.762$ ,  $p_{FDR} = .001$ ,  $d = -0.863$ , and EXT<sub>S</sub>,  $t(18) = -6.475$ ,  $p_{FDR} < .001$ ,  $d = -1.486$ . The EXT<sub>I</sub> amplitude, however, was significantly smaller than the one for EXT<sub>S</sub>,  $t(18) = -4.917$ ,  $p_{FDR} < .001$ ,  $d = -1.128$ .

Differences between performers and observers in the P2 component associated with the processing of EXT<sub>I</sub> were evident from visual inspection also with respect to the latency (see Figure 2A). A  $3 \times 2$  (Sound  $\times$  Group) ANOVA on the P2 latencies indeed revealed a significant Sound  $\times$  Group interaction,  $F(2, 72) = 9.574$ ,  $p < .001$ ,  $\eta_p^2 = .210$ . While the latencies of the P2 associated with the processing of ACT and EXT<sub>S</sub> did not differ between performers and observers (both  $p_{FDR} \geq .291$ ), the P2 latency for EXT<sub>I</sub> was significantly higher in performers than in observers (mean difference = 44.421 msec, 95% CI [25.167, 63.675]),  $t(36) = 4.679$ ,  $p_{FDR} < .001$ ,  $d = 1.518$ .

### Preaction Potential

The group-averaged preaction potentials related to left and right button press at the C3 and C4 electrode sites in both experimental groups are shown in Figure 3. A  $2 \times 2 \times 2$  (Hand  $\times$  Hemisphere  $\times$  Group) ANOVA revealed a significant main effect of Group,  $F(1, 36) = 12.731$ ,  $p = .001$ ,  $\eta_p^2 = .261$ , indicating a greater negativity before the onset of button presses in the performers ( $M = -2.268$   $\mu$ V, 95% CI [-3.162, -1.375]) compared to the observers ( $M = -0.045$   $\mu$ V, 95% CI [-0.939, 0.848]). The Hand  $\times$  Hemisphere interaction was significant,  $F(1, 36) = 17.235$ ,  $p < .001$ ,  $\eta_p^2 = .324$ , as well as the Hand  $\times$  Hemisphere  $\times$  Group interaction,  $F(1, 36) = 16.657$ ,  $p < .001$ ,  $\eta_p^2 = .316$ . We further examined the three-way interaction by applying a  $2 \times 2$  (Hand  $\times$  Hemisphere) ANOVA separately for each group. For the performers, a significant Hand  $\times$  Hemisphere interaction was found,  $F(1, 18) = 23.065$ ,  $p < .001$ ,  $\eta_p^2 = .562$ . As expected, we observed more negative preaction potential amplitudes in the right than in the left hemisphere

for button presses performed with the left hand,  $t(18) = 2.669$ ,  $p_{FDR} = .016$ ,  $d = 0.612$ . Vice versa, more negative preaction potential amplitudes were found in the left than in the right hemisphere for button presses performed with the right hand,  $t(18) = -3.894$ ,  $p_{FDR} = .002$ ,  $d = -0.893$ . For the observers, neither the main effects nor the Hand  $\times$  Hemisphere interaction effect was significant (all  $p \geq .112$ ).

### Correlation Analyses

For the performers, analyses showed that the correlation (Pearson's  $r$ ) between the mean preaction potential and the N1 amplitude reduction for ACT computed with respect to either EXT<sub>I</sub> or EXT<sub>S</sub> was not significant at FCz (both  $p_{FDR} \geq .692$ ). Similarly, no significant correlation was observed for the observers (both  $p_{FDR} \geq .192$ ). Furthermore, we did not find any significant correlation between the variability of the timing of the observed actions and the N1 amplitude as well as the N1 reductions for the observers (all  $p_{FDR} \geq .486$ ).

### DISCUSSION

By simultaneously recording EEG data from two participants, we provided novel evidence for a differential modulation of ERP components associated with the sensory processing of auditory outcomes of performed versus observed actions. For performers, we replicated previous findings that action-generated sounds elicited an N1 and a P2 of reduced amplitude compared to externally generated sounds at FCz. For observers, we present first data that, when timing of the observed action is unpredictable (differently from Ghio et al., 2018), the N1 amplitude for sounds generated by observed button presses was not attenuated relative to the N1 for externally generated sounds. Crucially, however, also for observers, we found a reduction of the later P2 amplitude for processing observed action versus externally generated sounds. Our study thus contributes to the extant research by showing a different modulation pattern of the N1 and P2 components in action observation versus a similar attenuation of both components in action performance for processing action-generated sounds.

### Differences between Performers and Observers in the Early Phase of Auditory Action Outcome Processing

Differences between action performance and observation in the early phase of auditory action outcome processing as indicated by our results for the N1 component can be interpreted by taking into account one class of studies focusing on the role of cerebellar forward model predictions, which suggest a functional dissociation between the N1 and the P2 component. In an EEG study on cerebellar lesion

patients, Knolle et al. (2013) found a reduction of the P2 amplitude for processing self versus externally generated sounds both in patients and healthy controls, while only the latter showed also a reduction of the N1 amplitude. A magnetoencephalography study by Cao et al. (2017) showed that processing action-generated sounds (delayed by 100 msec) induced an amplitude attenuation of both the mN1, which originated from the auditory cortex (see also Aliu, Houde, & Nagarajan, 2009; Martikainen, Kaneko, & Hari, 2005), and the mP2, whose source was localized in the inferior frontal gyrus/insula. The findings also revealed a stronger mN1 but not mP2 attenuation in the adaptation to the delay between the action and the sound, which was abolished by cerebellar transcranial magnetic stimulation (Cao et al., 2017). These studies suggest that the modulation of the N1 component reflects motor-based cerebellar predictions, possibly calculated by forward models based on efference copies. Although for environment-related action outcomes, such as the tones of this study, predictions can also be based on other, nonmotoric types of information (in line with predictive coding accounts), early processing in the N1 time window can only be affected by concomitant actions, when action-related information is available before movement onset. Our interpretation of the selective N1 reduction in response to action-generated sounds for performers is thus that an efference copy is either not generated at all in action observation or too late to affect the N1, at least when the timing of the observed action is unpredictable. These findings thus also extend and clarify the result pattern we previously obtained (Ghio et al., 2018), and highlight the importance of the timing of prediction generation in action observation.

### **Similarities between Performers and Observers in the Later Phase of Auditory Action Outcome Processing**

Concerning the later positive ERP component, there are inconsistencies across studies with respect to its latency and this component has consequently been labeled as either P2 or P3a. For its interpretation, we focus on another class of studies reporting a functional dissociation between N1 and P2/P3a and the resulting two main interpretations. First, it has been suggested that the P2/P3a reflects stimulus (un)predictability: The lower the predictability, the higher the amplitude (Neszmélyi & Horváth, 2017; Knolle et al., 2013; Baess et al., 2011). Contrary to the N1, however, the P2/P3a amplitude reduction does not seem to reflect motor-based, cerebellar-dependent predictions as it is independent from the disruption of cerebellar information processing (Cao et al., 2017; Knolle et al., 2013). In turn, a P2 but not an N1 amplitude reduction was found for sounds that were only visually cued (Sowman, Kuusik, & Johnson, 2012). These studies suggest that a modulation of the P2/P3a amplitude can reflect motor-independent predictive mechanisms in accordance with predictive coding accounts.

Our findings of a reduction of the positive component for processing action- versus externally generated sounds both in action performance and observation can thus be indicative of motor-independent predictive mechanisms in both conditions at this later stage. It should be noted, however, that we found a larger and later amplitude peak for processing intermixed externally generated sounds in action performance, which, based on post hoc examinations, could not be because of a contamination of the analyzed segments by button presses or because of pooling between different EXT<sub>1</sub>, because the waveforms were comparable for the different EXT<sub>1</sub>. Such larger and later amplitude peaks for EXT<sub>1</sub> only in performers might indicate a lower predictability for these sounds within a context in which other sounds can be predicted as a result of self-performed actions (similar findings in Knolle et al., 2013). Speculatively, this finding might suggest that motor-dependent and motor-independent prediction mechanisms might interact in action performance in a different way compared to action observation (see also Dogge et al., 2019).

The P2/P3a attenuation in observers, however, does not necessarily hint at motor-independent prediction mechanisms. As outlined in the Introduction, the involvement of premotor regions, which are a likely source of efference copies (Timm et al., 2014; Waszak et al., 2012) and responsible for sensory attenuation (Voss, Ingram, Haggard, & Wolpert, 2006), occurs later for action observation compared to performance. In a magnetoencephalography study, Nishitani and Hari (2000) investigated the cortical activation sequence in active versus observed pinch movements: Whereas in performance, activation starts from the premotor areas and extends into the motor cortex, in action observation, it starts in the visual cortex, then extends into the premotor areas, and then to the motor cortex (see also Sebastiani et al., 2014). The activation of the premotor areas occurred thus about 100 msec later in action observation (Nishitani & Hari, 2000). In observers, sensory prediction may thus be based on motor simulation after the observed action, which occurs too late to affect early processing in the N1 time window, but in time for later processing in the P2/P3a time window. This may be different for observed actions whose onset time is perfectly predictable, for which motor preparation processes are similar as for performed actions, as revealed by the preaction/readiness potential (Kilner, Vargas, Duval, Blakemore, & Sirigu, 2004). In particular, in the study by Kilner et al. (2004), the onset of the action was always predictable based on a visual cue. In this study, in turn, the exact observed actions' onset times were unpredictable. In this sense, the fact that we did not find a significant preaction potential in observers, differently from Kilner et al. (2004), might suggest that preaction potential in action observation, as well as motor-based predictions, might crucially depend on action's onset predictability. In future studies, additional control conditions might help to clarify the relative contribution of motor- and non-motor-based predictive mechanisms in action observation and the timing of such modulations.

## Implications of Similarities and Differences between Performers and Observers in Auditory Action Outcome Processing for Agency Attribution

A different modulation of N1 and P2/P3a in action performance and observation as found in this study might provide important insights for the debate about the sense of agency for action outcomes. According to both sensorimotor control (Wolpert & Flanagan, 2001) and predictive coding accounts (Brown, Adams, Pares, Edwards, & Friston, 2013; Friston, Mattout, & Kilner, 2011), sensory attenuation contributes to the sense of agency as the attenuated sensory stimulus can be recognized as self-generated (Reznik & Mukamel, 2019; Picard & Friston, 2014; Blakemore & Frith, 2003; Blakemore, Wolpert, & Frith, 2002). Based on previous evidence and according to the dual step account of agency (Synofzik, Vosgerau, & Voss, 2013; Synofzik, Vosgerau, & Newen, 2008), we can speculate that a modulation of the N1 and the P2 amplitude might reflect two aspects of the sense of agency (see also Ghio et al., 2018; Knolle et al., 2013). The N1 attenuation might contribute to the feeling of self-agency, that is, a first-person, implicit feeling that the action was caused by yourself (Synofzik et al., 2008). In turn, the P2/P3a attenuation might contribute to agency judgment, that is, the explicit attribution of agency either to the self or to another agent (Synofzik et al., 2008). This speculation is based on studies in which the sense of agency for self-generated stimuli was contextually manipulated by either inducing an illusory sense of agency (Timm, Schönwiesner, Schröger, & SanMiguel, 2016) or by creating uncertainties in the sense of agency (Kühn et al., 2011; for behavioral data, see Desantis, Weiss, Schütz-Bosbach, & Waszak, 2012). Timm et al. (2016) showed that the perception of agency correlated with a modulation of P2 but not N1 effects. Kühn et al. (2011) found a greater reduction of the P3a, but not of the N1 amplitude, for sounds judged as self- versus externally generated. Further studies are needed to test our speculative interpretation by measuring indices of agency feeling and attribution both in action performance and observation.

## Conclusion

This study shows time-dependent similarities and differences in the influence of performed and observed actions on the processing of sensory action outcomes. Early processing (N1) of action-generated sounds is altered only by self-performed actions, probably because (efference copy-based) motor information is not available early enough in the observation of actions with unpredictable onset timing. In turn, later sound processing (P2) was altered not only by self-performed actions but also by observed actions. In how far the modulation by observed actions reflects motor-dependent or independent prediction mechanisms and to what extent it depends on the predictability of the observed movement need to be examined in future studies.

Reprint requests should be sent to Marta Ghio, CIMEC - Center for Mind/Brain Sciences, University of Trento, Via delle Regole 101, 38123 Mattarello (TN), Italy, or via e-mail: martavirginia.ghio@unitn.it.

## Diversity in Citation Practices

A retrospective analysis of the citations in every article published in this journal from 2010 to 2020 has revealed a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were M(an)/M = .408, W(oman)/M = .335, M/W = .108, and W/W = .149, the comparable proportions for the articles that these authorship teams cited were M/M = .579, W/M = .243, M/W = .102, and W/W = .076 (Fulvio et al., *JoCN*, 33:1, pp. 3–7). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

## REFERENCES

- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, 11, 1109–1116. DOI: <https://doi.org/10.1038/nn.2182>, PMID: 19160510
- Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience*, 21, 791–802. DOI: <https://doi.org/10.1162/jocn.2009.21055>, PMID: 18593265, PMCID: PMC2944400
- Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: Past, present and future. *Neuroscience & Biobehavioral Reviews*, 44, 76–93. DOI: <https://doi.org/10.1016/j.neubiorev.2012.07.006>, PMID: 22917915, PMCID: PMC3522775
- Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*, 48, 1276–1283. DOI: <https://doi.org/10.1111/j.1469-8986.2011.01196.x>, PMID: 21449953
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57, 289–300. DOI: <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Blakemore, S.-J., & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, 13, 219–224. DOI: [https://doi.org/10.1016/S0959-4388\(03\)00043-6](https://doi.org/10.1016/S0959-4388(03)00043-6), PMID: 12744977
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, 6, 237–242. DOI: [https://doi.org/10.1016/S1364-6613\(02\)01907-1](https://doi.org/10.1016/S1364-6613(02)01907-1), PMID: 12039604
- Brown, H., Adams, R. A., Pares, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, 14, 411–427. DOI: <https://doi.org/10.1007/s10339-013-0571-3>, PMID: 23744445, PMCID: PMC3824582
- Brown, E. C., & Brüne, M. (2012). Evolution of social predictive brains? *Frontiers in Psychology*, 3, 414. DOI: <https://doi.org/10.3389/fpsyg.2012.00414>, PMID: 23181028, PMCID: PMC3501286

- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, *16*, 1905–1910. **DOI:** <https://doi.org/10.1016/j.cub.2006.07.065>, **PMID:** 17027486
- Cao, L., & Gross, J. (2015). Cultural differences in perceiving sounds generated by others: Self matters. *Frontiers in Psychology*, *6*, 1865. **DOI:** <https://doi.org/10.3389/fpsyg.2015.01865>, **PMID:** 26696931, **PMCID:** PMC4667006
- Cao, L., Veniero, D., Thut, G., & Gross, J. (2017). Role of the cerebellum in adaptation to delayed action effects. *Current Biology*, *27*, 2442–2451. **DOI:** <https://doi.org/10.1016/j.cub.2017.06.074>, **PMID:** 28781049, **PMCID:** PMC5571438
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, *9*, 587–600. **DOI:** <https://doi.org/10.1038/nrn2457>, **PMID:** 18641666, **PMCID:** PMC5153363
- Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4–14. **DOI:** <https://doi.org/10.1016/j.brainres.2005.12.115>, **PMID:** 16460715
- Desantis, A., Weiss, C., Schütz-Bosbach, S., & Waszak, F. (2012). Believing and perceiving: Authorship belief modulates sensory attenuation. *PLoS One*, *7*, e37959. **DOI:** <https://doi.org/10.1371/journal.pone.0037959>, **PMID:** 22666424, **PMCID:** PMC3362539
- Dogge, M., Custers, R., & Aarts, H. (2019). Moving forward: On the limits of motor-based forward models. *Trends in Cognitive Sciences*, *23*, 743–753. **DOI:** <https://doi.org/10.1016/j.tics.2019.06.008>, **PMID:** 31371239
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. **DOI:** <https://doi.org/10.3758/BF03193146>, **PMID:** 17695343
- Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*, 769–771. **DOI:** <https://doi.org/10.1038/nature01861>, **PMID:** 12917683
- Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2014). Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone. *Schizophrenia Bulletin*, *40*, 804–812. **DOI:** <https://doi.org/10.1093/schbul/sbt072>, **PMID:** 23754836, **PMCID:** PMC4059422
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*, 137–160. **DOI:** <https://doi.org/10.1007/s00422-011-0424-z>, **PMID:** 21327826, **PMCID:** PMC3491875
- Ghio, M., Scharmach, K., & Bellebaum, C. (2018). ERP correlates of processing the auditory consequences of own versus observed actions. *Psychophysiology*, *55*, e13048. **DOI:** <https://doi.org/10.1111/psyp.13048>, **PMID:** 29266338
- Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54–65. **DOI:** <https://doi.org/10.1016/j.brainres.2015.03.038>, **PMID:** 25843932
- JASP Team. (2019). JASP (Version 0.11.1) [Computer software]. Retrieved from <https://jasp-stats.org/>.
- Keyesers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition. *Trends in Cognitive Sciences*, *11*, 194–196. **DOI:** <https://doi.org/10.1016/j.tics.2007.02.002>, **PMID:** 17344090
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8*, 159–166. **DOI:** <https://doi.org/10.1007/s10339-007-0170-2>, **PMID:** 17429704, **PMCID:** PMC2649419
- Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, *7*, 1299–1301. **DOI:** <https://doi.org/10.1038/nn1355>, **PMID:** 15558063
- Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*, 2449–2461. **DOI:** <https://doi.org/10.1016/j.cortex.2012.12.012>, **PMID:** 23318086
- Kühn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., & Voss, M. (2011). Whodunnit? Electrophysiological correlates of agency judgements. *PLoS One*, *6*, e28657. **DOI:** <https://doi.org/10.1371/journal.pone.0028657>, **PMID:** 22194878, **PMCID:** PMC3237473
- Lange, K. (2011). The reduced N1 to self-generated tones: An effect of temporal predictability? *Psychophysiology*, *48*, 1088–1095. **DOI:** <https://doi.org/10.1111/j.1469-8986.2010.01174.x>, **PMID:** 21261634
- Martikainen, M. H., Kaneko, K.-I., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*, 299–302. **DOI:** <https://doi.org/10.1093/cercor/bhh131>, **PMID:** 15238430
- Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, *14*, 2135–2137. **DOI:** <https://doi.org/10.1097/00001756-200312020-00001>, **PMID:** 14625435
- Neszmélyi, B., & Horváth, J. (2017). Consequences matter: Self-induced tones are used as feedback to optimize tone-eliciting actions. *Psychophysiology*, *54*, 904–915. **DOI:** <https://doi.org/10.1111/psyp.12845>, **PMID:** 28240775
- Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences, U.S.A.*, *97*, 913–918. **DOI:** <https://doi.org/10.1073/pnas.97.2.913>, **PMID:** 10639179, **PMCID:** PMC15430
- Picard, F., & Friston, K. (2014). Predictions, perception, and a sense of self. *Neurology*, *83*, 1112–1118. **DOI:** <https://doi.org/10.1212/WNL.0000000000000798>, **PMID:** 25128179, **PMCID:** PMC4166359
- Pickering, M. J., & Clark, A. (2014). Getting ahead: Forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, *18*, 451–456. **DOI:** <https://doi.org/10.1016/j.tics.2014.05.006>, **PMID:** 24909775
- Poonian, S. K., McFadyen, J., Ogden, J., & Cunnington, R. (2015). Implicit agency in observed actions: Evidence for N1 suppression of tones caused by self-made and observed actions. *Journal of Cognitive Neuroscience*, *27*, 752–764. **DOI:** [https://doi.org/10.1162/jocn\\_a\\_00745](https://doi.org/10.1162/jocn_a_00745), **PMID:** 25321488
- Reznik, D., & Mukamel, R. (2019). Motor output, neural states and auditory perception. *Neuroscience & Biobehavioral Reviews*, *96*, 116–126. **DOI:** <https://doi.org/10.1016/j.neubiorev.2018.10.021>, **PMID:** 30391407
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274. **DOI:** <https://doi.org/10.1038/nrn2805>, **PMID:** 20216547
- Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of Neurophysiology*, *96*, 1358–1369. **DOI:** <https://doi.org/10.1152/jn.00227.2006>, **PMID:** 16687620
- Sato, A. (2008). Action observation modulates auditory perception of the consequence of others' actions. *Consciousness and Cognition*, *17*, 1219–1227. **DOI:** <https://doi.org/10.1016/j.concog.2008.01.003>, **PMID:** 18299207
- 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. **DOI:** <https://doi.org/10.1111/ejn.12816>, **PMID:** 25728182, **PMCID:** PMC4402002

- Sebastiani, V., de Pasquale, F., Costantini, M., Mantini, D., Pizzella, V., Romani, G. L., et al. (2014). Being an agent or an observer: Different spectral dynamics revealed by MEG. *Neuroimage*, *102*, 717–728. **DOI:** <https://doi.org/10.1016/j.neuroimage.2014.08.031>, **PMID:** 25175536
- Sowman, P. F., Kuusik, A., & Johnson, B. W. (2012). Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Experimental Brain Research*, *222*, 149–157. **DOI:** <https://doi.org/10.1007/s00221-012-3204-7>, **PMID:** 22885999, **PMCID:** PMC3898150
- Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*, 219–239. **DOI:** <https://doi.org/10.1016/j.concog.2007.03.010>, **PMID:** 17482480
- Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: An interplay between prediction and postdiction. *Frontiers in Psychology*, *4*, 127. **DOI:** <https://doi.org/10.3389/fpsyg.2013.00127>, **PMID:** 23508565, **PMCID:** PMC3597983
- 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. **DOI:** [https://doi.org/10.1162/jocn\\_a\\_00552](https://doi.org/10.1162/jocn_a_00552), **PMID:** 24392902
- Timm, J., Schönwiesner, M., Schröger, E., & SanMiguel, I. (2016). Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. *Cortex*, *80*, 5–20. **DOI:** <https://doi.org/10.1016/j.cortex.2016.03.018>, **PMID:** 27137101
- Voss, M., Ingram, J. N., Haggard, P., & Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature Neuroscience*, *9*, 26–27. **DOI:** <https://doi.org/10.1038/nn1592>, **PMID:** 16311591, **PMCID:** PMC2636578
- Waszak, F., Cardoso-Leite, P., & Hughes, G. (2012). Action effect anticipation: Neurophysiological basis and functional consequences. *Neuroscience & Biobehavioral Reviews*, *36*, 943–959. **DOI:** <https://doi.org/10.1016/j.neubiorev.2011.11.004>, **PMID:** 22108008
- Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011a). The self in social interactions: Sensory attenuation of auditory action effects is stronger in interactions with others. *PLoS One*, *6*, e22723. **DOI:** <https://doi.org/10.1371/journal.pone.0022723>, **PMID:** 21818373, **PMCID:** PMC3144943
- Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011b). The self in action effects: Selective attenuation of self-generated sounds. *Cognition*, *121*, 207–218. **DOI:** <https://doi.org/10.1016/j.cognition.2011.06.011>, **PMID:** 21784422
- Weiss, C., & Schütz-Bosbach, S. (2012). Vicarious action preparation does not result in sensory attenuation of auditory action effects. *Consciousness and Cognition*, *21*, 1654–1661. **DOI:** <https://doi.org/10.1016/j.concog.2012.08.010>, **PMID:** 23069246
- Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, *12*, 739–751. **DOI:** <https://doi.org/10.1038/nrn3112>, **PMID:** 22033537
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *358*, 593–602. **DOI:** <https://doi.org/10.1098/rstb.2002.1238>, **PMID:** 12689384, **PMCID:** PMC1693134
- Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*, R729–R732. **DOI:** [https://doi.org/10.1016/S0960-9822\(01\)00432-8](https://doi.org/10.1016/S0960-9822(01)00432-8), **PMID:** 11566114
- Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*, 1212–1217. **DOI:** <https://doi.org/10.1038/81497>, **PMID:** 11127840