Comparing the Processing of Auditory Action Consequences in Action Performance and Observation

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Sophie Egan
aus Bingen am Rhein

Düsseldorf, August 2025

aus dem Institut für Experimentelle Psychologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Professor Dr. Christian Bellebaum

2. Professor Dr. Eckart Zimmermann

Tag der mündlichen Prüfung: 03.11.2025

Table of Contents

List of Abbreviations

AICS	Auckland individualism collectivism scale
ANOVA	Analysis of variance
AQ	Anthropomorphism questionnaire
CAPS	Cardiff anomalous perceptions scale
EEG	Electroencephalography
EQ	Empathy quotient
ERP	Event-related potential
fMRI	Functional magnetic resonance imaging
GAToRS	General attitudes towards robots scale
JND	Just noticeable difference
MEG	Magnetoencephalography
PDI	Peters et al. delusions inventory
PSE	Point of subjective equality
SCS	Self-construal scale
TMS	Transcranial magnetic stimulation
VR	Virtual reality

List of Figures

Abstract

Sensory attenuation describes the phenomenon of reduced sensory processing for self-generated input compared to externally generated input (e.g., perception as less loud). A typical measurement of sensory attenuation is through neurophysiological response, in particular N1 and P2 components in electroencephalography (EEG). Predictive mechanisms were proposed to underlie sensory attenuation. These include motor-based cerebellar forward models as well as general predictive mechanisms. A function of sensory attenuation may be the distinction between own and external effects, and thus a contribution to the sense of agency. Due to similar neural activation during action observation and performance, the same mechanisms could be applied for outcomes of observed actions and yield similar attenuation relative to other external input. Yet, perceptual and neurophysiological measures have shown inconsistent findings regarding sensory attenuation in action observation. In this dissertation, sensory attenuation in action performance and observation was further investigated in three EEG studies to assess potential effects of temporal predictability, culture, and agency as well as to examine the role of prediction errors, which might help explain previous inconsistencies and clarify the underlying mechanisms. A novel addition in these studies compared to previous studies was the inclusion of the cue condition with temporally predictable external tones. In the first study, we found the N1 to reflect temporal predictability and the P2 attenuation in action performance and observation to go beyond it, hinting at further factors such as the sense of agency. We further found no indication of an influence of culture on attenuation in action observation, which studies on perceptual attenuation had suggested. In the second study, to explore effects of agency attribution we extended action-observation conditions to include a robotic hand controlled by either a human agent or the computer. However, the P2 did not reflect differences in agency attribution. Instead, there was a self-specific P2 attenuation beyond the attenuation for temporally predictable tones. The type of action-observation setting, which was live in the first study and via video in the second study, might modulate the mechanisms applied in action observation. The third study focused on the modulation of N1 and P2 amplitudes by prediction errors resulting from tones occurring in only half of the trials. For this purpose, we used reinforcement learning models to calculate the prediction error in each trial. In this context, correct predictions or enacting control could be perceived as rewarding. N1 amplitudes were smaller when the calculated prediction error was smaller. This effect of prediction error was stronger for self-generated tones. Smaller amplitudes with smaller errors and a stronger effect for self-generation were also found for the P2, but overall, the P2 was only attenuated for action observation and cueing, not self-generation, relative to external tones. Thus, both components showed self-specific effects relative to the video action-observation and cue conditions, and only the P2 was sensitive to the lack of a contingent relationship between action and tone. In sum, processing auditory consequences in action performance and observation appeared to rely on different mechanisms particularly when the observation was implemented via video. Moreover, the N1 and P2 components appeared to be functionally distinct, in line with previous studies, in that the former likely relates to temporal predictability and cerebellar forward models, whereas the latter could reflect the identification of the tone as an outcome of an event, particularly self-performed action. Future research might further examine the relationship of perceptual and neurophysiological measures, the modulations of action-observation setting and context, and disentangle the different mechanisms involved in what we measure as sensory attenuation.

1 General Introduction

1.1 The Sensory Attenuation Phenomenon

When typing on your keyboard, you might not take much notice of the keyboard clicks you produce, but someone else in the same room might find the continuous stream of clicks rather annoying. Indeed, there is evidence that self-generated sensory input is processed differently than externally generated input and one phenomenon regarding this comparison is sensory attenuation (e.g., Korka et al., 2022). Sensory attenuation is typically defined as a reduction of sensory processing of the self-generated sensory input compared to externally generated input and, although it has been subject of much research, the underlying mechanisms are not fully understood (for reviews, see Horváth, 2015; Hughes et al., 2013; Korka et al., 2022). The following sub-sections give an overview of how sensory attenuation has been measured in previous studies, what mechanisms have been proposed to lead to sensory attenuation, and of the relationship between sensory attenuation and the sense of agency.

1.1.1 Measures of Sensory Attenuation

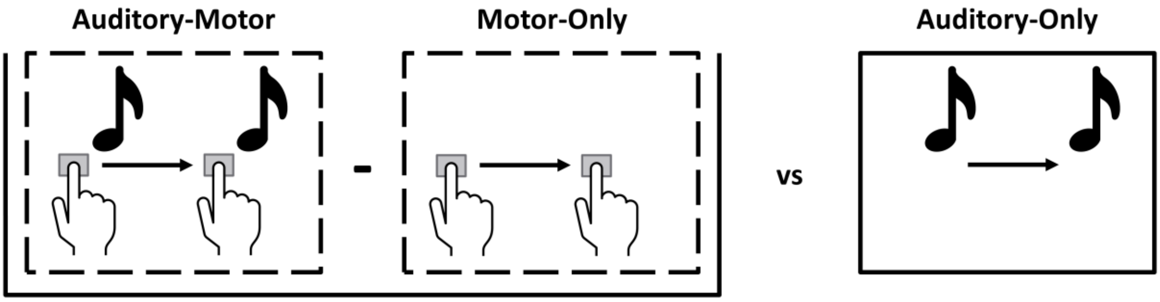
In the first studies measuring the phenomenon of sensory attenuation, stimuli in the somatosensory modality (Weiskrantz et al., 1971) and the auditory as well as the visual modality (Schafer & Marcus, 1973) were examined. The former addressed the question of why one cannot tickle oneself with ticklishness-comparison judgements, while the latter employed electroencephalography (EEG) to compare components of event-related potentials (ERPs) as a more direct measure of the brain's response. Following research has distinguished between auditory, visual, and somatosensory modalities; and perceptual (i.e., behavioural) and neurophysiological (i.e., typically with ERP components) measures of sensory attenuation.

The auditory stimuli in the initial studies were clicks (Schafer & Marcus, 1973) or pure tones (McCarthy & Donchin, 1976) that were generated by self-timed button presses or externally. This approach (particularly with tones) has been widely applied in following studies (Horváth, 2015), but some studies on auditory attenuation have investigated self-generated and played-back speech (e.g., Curio et al., 2000; Ford et al., 2001). However, attenuation for self-generated speech (i.e., vocalisation) involves peripheral confounds and may rather be related to specialised mechanisms (Horváth, 2015; but see Mifsud & Whitford, 2017). A differentiation of underlying mechanisms was also proposed as speech and tones were categorised as two different types of action outcomes (Dogge, Custers, et al., 2019). In the following, auditory attenuation refers to attenuation of tone processing unless otherwise specified.

Perceptual measures of auditory attenuation have been obtained with a comparison task (e.g., Sato, 2008). A standard tone with fixed loudness was either self-generated or externally generated. This was followed by a comparison tone with variable loudness (but see Kiepe et al., 2024, for switched loudness properties) and participants were tasked to determine which tone was louder. Based on the responses, the point of subjective equality (PSE) and the just noticeable difference (JND) can be calculated. Typically, sensory attenuation was defined as a reduced PSE of self-generated tones compared to the PSE of externally generated ones, while the JND reflects a general ability to differentiate loudness levels (e.g., Weiss et al., 2011a). The lack of sensory attenuation in JND values has been interpreted as intact sensitivity for self-generated tones, and thus, attenuation as reflecting the systematic reduction of intensity (e.g., Weiss et al., 2011a). An alternative approach to measuring auditory attenuation is the application of an auditory detection task. However, with this approach an enhancement of sensitivity (i.e., better detection) rather than attenuation was found for self-generated tones (e.g., Reznik et al., 2014). These findings were reconciled by the study of Reznik et al. (2015) that applied a modified comparison task with near- or supra-threshold tones. For self-generated tones compared to cued tones, they showed enhancement with near-threshold and attenuation with supra-threshold intensities. Including both types of task, these findings were extended, showing the same interaction effect on PSE (i.e., perceptual bias) but no effect on sensitivity in the comparison (via JND) or the detection task (Paraskevoudi & SanMiguel, 2021).

As mentioned above, neurophysiological studies on auditory attenuation have widely applied a similar approach as Schafer and Marcus (1973) and McCarthy and Donchin (1976). Referred to as the contingent paradigm (Horváth, 2015), it encompasses three types of experimental blocks: (i) an auditory-motor, (ii) an auditory-only, and (iii) a motor-only block. The tasks of these blocks are typically (i) generating tones with a self-timed action (e.g., button press), (ii) passively listening to externally generated tones, and (iii) performing the self-timed action without the presentation of tones (see Figure 1.1.1). Tone processing is typically measured via the auditory N1 and P2 ERP components or their counterparts in magnetoencephalography (MEG; Horváth, 2015; Korka et al., 2022). These ERPs of the auditory-motor block are compared to those of the auditory-only block to assess attenuation but because of the activity related to the generation itself, the difference of the auditory-motor and motor-only block is used for the comparison. The calculation of the difference, termed motor correction, has been criticised because of the differences in ERPs between actions with and without sensory effects (Horváth, 2015; Neszmélyi et al., 2022) but is yet widely applied (e.g., Klaffehn et al., 2019). Other criticisms of this paradigm relate to confounds in the comparison of self-generated and external tones: predictability of the temporal onset (i.e., when-prediction), control over the temporal onset (i.e., temporal control), and predictability of stimulus identity (i.e., tone frequency; what-prediction) based on motor and non-motor cues (Hughes et al., 2013).

FIGURE 1.1.1: VISUALISATION OF THE CONTINGENT PARADIGM



Note. The auditory-motor block is also referred to as self-generation block, the motor-only block as control block for self-generation, the corrected auditory-motor as self-generation condition, and the auditory-only block as external-tone block or condition.

Of the two ERP components, the focus has been on the N1 (Horváth, 2015), though its sub-components (Näätänen & Picton, 1987) are not typically differentiated (for an example of an anylsis of sub-components, see SanMiguel, Todd, et al., 2013). Defined as the negative peak occurring 50 to 150 ms after tone onset (Näätänen & Picton, 1987) at frontocentral electrodes, the amplitudes of the N1 were found to be reduced for self-generated compared to the external tones (e.g., Klaffehn et al., 2019). The P2, a subsequent positivity peaking 150 to 250 ms after tone onset at central electrodes (Crowley & Colrain, 2004), has been observed to also be attenuated for self-generated compared to external tones (e.g., Klaffehn et al., 2019). Both components have been found to increase in amplitudes, for instance, with increasing tone loudness or inter-tone intervals (Crowley & Colrain, 2004; Woods, 1995). Indeed, in some studies on sensory attenuation the components are analysed as an N1-P2 complex (e.g., Timm et al., 2014). However, there is evidence for treating the N1 and P2 as distinct components (Crowley & Colrain, 2004; Horváth, 2015). In one study on sensory attenuation, for example, the N1 but not the P2 of self-generated tones was attenuated compared to external tones when an action-outcome delay was introduced (Klaffehn et al., 2019).

Regarding the relationship between perceptual and neurophysiological measures of auditory attenuation, the two studies investigating this did not find attenuation in either perceptual or neurophysiological measures (Dogge, Hofman, et al., 2019; Ody et al., 2023). This might be due to the modifications of the paradigm they applied with either voluntary versus involuntary action (Ody et al., 2023) or congruent versus incongruent action outcomes (Dogge, Hofman, et al., 2019). Yet, overall, perceptual and neurophysiological measures roughly correspond in showing auditory attenuation of self-generated stimuli compared to externally generated stimuli.

1.1.2 Mechanisms Underlying Sensory Attenuation

To explain the attenuation of processing of self-generated sensory input compared to externally generated input, various mechanisms have been proposed and discussed (e.g., Horváth, 2015). A common understanding is that because the self-generated sensory input is more predictable than the external one, the processing is less relevant, but how the predictions are formed and how this leads to attenuation is under debate (e.g., Dogge, Custers, et al., 2019).

Sensory attenuation has been mostly interpreted in terms of internal forward models (Blakemore, Wolpert, et al., 2000; Wolpert & Flanagan, 2001). Theories of forward models state that a copy of the motor command, termed the efference copy, is sent to the forward model, which derives predictions of the sensory input generated by the action. The predicted and perceived sensory input are then compared. The more the prediction and the actual input match, the more attenuation is observed. According to one account, the correctly predicted input cancels out the actual input (e.g., Wolpert & Flanagan, 2001), but alternative accounts have been discussed (e.g., pre-activation, see Dogge, Custers, et al., 2019; Horváth, 2015; Roussel et al., 2013). It was further proposed that the forward model is represented in the cerebellum (e.g., Blakemore, Wolpert, et al., 2000; see also Shadmehr et al., 2010), which is in line with results of studies applying functional magnetic resonance imaging (fMRI) to examine sensory attenuation in various modalities (Blakemore et al., 1998; Kilteni & Ehrsson, 2020; Straube et al., 2017). Interestingly, in the somatosensory modality, perceptual attenuation correlated with the functional connectivity between the cerebellum and the secondary somatosensory cortex (Kilteni & Ehrsson, 2020). In the auditory modality, N1 but not P2 attenuation (of self-generated vs. external tones) was absent in patients with cerebellar lesions (Knolle et al., 2013a). Furthermore, the adaptation of the MEG N1 (i.e., M100) attenuation (i.e., increased attenuation over time) to an action-outcome delay of 100 ms was not observed after suppressive transcranial magnetic stimulation (TMS) of the cerebellum (Cao et al., 2017). Thus, N1 attenuation and its adaptation was shown to be dependent on the cerebellum.

Further support for forward models came from studies inducing involuntary movements (i.e., without a motor command) and the comparison of their outcomes to outcomes of voluntary actions: Attenuation relative to external tones was found for outcomes only of voluntary actions, not of movements induced by TMS applied to the primary motor cortex (neurophysiological attenuation in the auditory modality; Timm et al., 2014), of movements induced by stimulating the median nerve in the forearm (neurophysiological attenuation in auditory modality; Jack et al., 2021), or of the passive movement of letting the finger fall when the platform it rested on was removed (perceptual attenuation in somatosensory modality; Kilteni et al., 2020). Although these studies show the importance of a voluntary action for sensory attenuation, similar neurophysiological effects were found for (auditory) action outcomes and outcomes of in-action, that is, not performing an action in a given timeframe (e.g., Han et al., 2021). This further highlighted the role of intention rather than the action itself (see also Korka et al., 2022).

Dogge, Custers, et al. (2019) argued that the often arbitrary environment-related action-outcome associations, such as a button press and a tone, are learned too quickly for motor-based mechanisms to be effective. These, they proposed, are rather supported by general predictive mechanisms, while more body-related action-outcome associations, such as self-touch and speech, rely on the motor-based forward model (Dogge, Custers, et al., 2019).

A critical difference between self-generated tones and external tones has been the ability to predict the temporal onset of the tone (i.e., when-prediction) only for self-generation, which cannot exclude general predictive mechanisms as underlying the attenuation effect (Hughes et al., 2013). Previous studies have presented action-outcome tones immediately after the action (e.g., Knolle et al., 2013a; Sato, 2008), with a fixed delay (e.g., Elijah et al., 2018; Klaffehn et al., 2019; Lange, 2011), with a variable delay (e.g., Bäß et al., 2008; Lange, 2011; Pinheiro et al., 2019), or with a delay incongruent to the learnt temporal relationship (e.g., Aliu et al., 2009; Elijah et al., 2016; Timm et al., 2016). Some showed that the amplitudes were smallest when the outcome onset was predictable (e.g., Pinheiro et al., 2019; Timm et al., 2016, for N1) with a deviation tolerance for learnt delays (e.g., Aliu et al., 2009, for M100; Elijah et al., 2016, for N1), others that temporal predictability did not affect the amplitudes (e.g., Bäß et al., 2008, for N1; Lange, 2011, for N1). Some studies made external tones temporally predictable by temporal cueing (i.e., playing tones in a fixed interval; Schafer & Marcus, 1973) or presenting a visual cue before tone onset, such as a loading bar filling up (Klaffehn et al., 2019). Comparisons of self-generated and temporally predictable external tones further showed mixed results: attenuation of self-generated compared to the cued external tones (e.g., Harrison et al., 2021, for P2; Klaffehn et al., 2019; Lange, 2011, for N1; Mifsud, Oestreich, et al., 2016, for N1), reduced attenuation of self-generated tones, that is, additional attenuation for cued external tones compared to unpredictable external tones (e.g., Mifsud, Oestreich, et al., 2016, for P2; Schafer & Marcus, 1973, for N1), and no differences in amplitudes of self-generated and cued external tones (e.g., Harrison et al., 2021, for N1; Schafer & Marcus, 1973, for P2). N1 attenuation compared to unpredictable external tones was also observed for external tones when cued (Kaiser & Schütz-Bosbach, 2018; see also Sowman et al., 2012, for P2 attenuation) and when temporal onset was predictable based on a preceding tone sequence (Lange, 2009). Hence, several studies have indicated the relevance of temporal predictability and thus general predictive mechanisms for neurophysiological attenuation of (self-generated) tones relative to unpredictable external tones.

Predictive coding (e.g., Rao & Ballard, 1999) or processing (see also integral forward model, Pickering & Clark, 2014) is the framework that is discussed for such general predictive mechanisms (e.g., Dogge, Custers, et al., 2019; Kiepe et al., 2021). According to this framework, perception results from the minimisation of surprise (i.e., prediction errors). Unlike the internal forward models described above, one common mechanism underlies action and prediction, as proprioceptive predictions would result in prediction errors and these would be resolved by action (active inference; Pickering & Clark, 2014). Sensory attenuation then arises from lowering precision and drawing away attention from the predicted sensory input, because instead the precision (i.e., estimated reliability of the prediction or confidence) for proprioception is increased so that movement can ensue (Brown et al., 2013; Pickering & Clark, 2014).

Both forward models and predictive coding accounts offer an explanation of the observed attenuation effects. However, perceived intensity enhancement of self-generated tones was found for tones at near-threshold level (e.g., Paraskevoudi & SanMiguel, 2021). The opposing process theory by Press et al. (2020) proposed that after an enhancement based on the predictability of the stimuli, predicted stimuli relative to stimuli that cause high surprise appear attenuated (in the time range of N1 and P2 components). This latter process depends on the informativeness of the stimuli and distinction from noise, and hence, is not activated for unpredictable near-threshold tones (Press et al., 2020).

In conclusion, motor-based forward models have been proposed to explain sensory attenuation, but doubts have been raised (e.g., Dogge, Custers, et al., 2019). Instead, the attenuation of the perceptual and neurophysiological responses to self-generated stimuli may result from a combination of motor-based and general predictive mechanisms (e.g., Korka et al., 2022), though the direction of the difference to external stimuli could depend on contextual factors (Press et al., 2020). To what extent each of these mechanisms contribute to sensory attenuation and what factors the magnitude of contribution depends on remain open questions.

1.1.3 Me vs. the World: Sensory Attenuation as Implicit Measure of Agency

According to forward models but not to predictive coding, sensory attenuation, that is, the successful match of predicted and perceived sensory input, poses a distinction to externally generated stimuli and thus enables the sense of agency (Kiepe et al., 2021). This describes the “sense that I am the one who is causing or generating an action” (Gallagher, 2000, p. 15) including action outcomes (Synofzik et al., 2008). For example, the self-produced keyboard clicks are predicted, attenuated, and thus identified as one’s own. External keyboard clicks are not predicted, not attenuated, and thus identified as not-one's-own. According to the two-step (Synofzik et al., 2008) or optimal-cue-integration (Synofzik et al., 2013) account of agency, the feeling and the judgement of agency can be differentiated. The former is an experience that results from a combination of weighted cues, including (but not limited to) comparison results of forward models. The latter, on the other hand, refers to the explicit attribution of agency and is an integration of further contextual cues and intentions. In ambiguous cases, when the feeling of agency is reduced, the judgement of agency is relied upon to find the best explanation (Synofzik et al., 2008). The weighting of the cues for feeling and judgement of agency might also depend on their situational reliability and affective factors (Synofzik et al., 2013).

It has been proposed that sensory attenuation represents an implicit measure of agency (Hughes et al., 2013). As explicit measures of agency, on the other hand, participants have been asked whether they were the agent of the tone allowing for yes-or-no responses (e.g., Timm et al., 2016), or participants rated the degree of their agency (e.g., Sato, 2008). Regarding this interpretation of sensory attenuation, studies have investigated patients with abnormalities in both the sense of agency and attenuation. In patients with schizophrenia, for instance, agency has been found to be misattributed, such in the case of auditory hallucinations, delusions of control, or thought insertions, which might be related to altered forward models (e.g., Feinberg, 1978; Franck et al., 2001). Studies comparing attenuation of self-generated stimuli to externally generated stimuli in such patients and control participants found reduced attenuation in patients in somatosensory (for perceptual measures, e.g., Blakemore, Smith, et al., 2000; Shergill et al., 2005) and auditory modalities, including vocalisation (for N1, e.g., Ford et al., 2007; Ford et al., 2001) and tones (for N1, e.g., Ford et al., 2014; Whitford et al., 2011). Kort et al. (2017) further found a similarly reduced N1 attenuation for vocalisation in patients and healthy participants who received infusions of ketamine, which induces schizophrenia-like symptoms. Moreover, scores relating to symptoms were negatively correlated with sensory attenuation in several studies: Scores of delusional ideation (PDI; Peters et al., 2004; Peters et al., 1999) of healthy participants with perceptual somatosensory attenuation (Teufel et al., 2010), PDI scores of healthy participants with perceptual auditory attenuation (Cao & Gross, 2015), scores of the schizotypal personality questionnaire (Raine, 1991) in healthy participants with N1 attenuation of vocalisations (Oestreich et al., 2015), severity scores of auditory hallucinations and percentages of agency misattribution in patients with (lacking) N1 attenuation of vocalisations (Heinks-Maldonado et al., 2007), and severity scores of hallucinations in patients with (lacking) attenuation in the secondary somatosensory cortex measured via fMRI (Shergill et al., 2014).

In healthy participants, the explicit rating of agency was reduced when discrepancies, such as altered stimulus identity and temporal delays, were applied and the participants were previously instructed that in some cases the experimenter elicits the tones (Sato & Yasuda, 2005). With a similar setup, Kühn et al. (2011) found agency-related attenuation (i.e., lower amplitudes for perceived self-agency than perceived other-agency) of P3a (370-390 ms) but not of N1 amplitudes. The use of delays to induce agency uncertainties was applied in further auditory attenuation studies: Implementing three different delays, with one as the most likely (70%), participants were induced to believe that tones were generated by the experimenter (as implied by the displayed name, though the tones were always self-generated) in a simultaneous cued generation task (Desantis et al., 2012). PSE values were reduced when participants believed in self-agency. Focusing on P2 amplitudes, Weller et al. (2017) found P2 enhancement, instead of attenuation, for delayed tones. Filling the delay between action and tone (i.e., visual animation for 2 s), however, increased the rated agency, while P2 amplitudes were not affected (Weller et al., 2017). In a different manipulation of delays, Timm et al. (2016) reduced the perceived agency (i.e., yes-or-no responses) by omitting the previously learnt delay (of 200 ms) between cued actions and tones. This correlated with P2 but not N1 attenuation (Timm et al., 2016). Furthermore, by inducing high or low illusions of control (i.e., learning 70% or 30% contingency between action and a specific tone), Seidel et al. (2021) again showed an agency-related effect on P2 (i.e., reduction for high compared to low perceived control over self-generated tones) but not N1 attenuation. Lastly, the findings of absent attenuation when movement was involuntary (e.g., elicited via TMS; Timm et al., 2014) are also in line with the hypothesis that attenuation represents an implicit agency measure. Including semi-voluntary actions, Jack et al. (2021) found a linear effect of volition on N1 attenuation, while the P2 was sensitive to the type of voluntary initiation of the involuntary movement (i.e., P2 attenuation for finger-on-finger press, but not for stimulator-induced press).

All in all, attenuation and the sense of agency appear to be related (e.g., Heinks-Maldonado et al., 2007) but diverging findings (e.g., Weller et al., 2017) and that prior belief about agency could affect attenuation (e.g., Desantis et al., 2012) indicate that this relation is not causal as forward models proposed. For neurophysiological auditory attenuation, N1 attenuation might result from cerebellar forward models contributing to the feeling of agency, while the later P2 attenuation rather reflects a more retrospective agency judgement (e.g., Knolle et al., 2013a; Seidel et al., 2021; Synofzik et al., 2013; Timm et al., 2016).

1.2 Me vs. You: Sensory Attenuation in Action Observation

As described above, your own keyboard clicks might escape your notice, while disturbing another person in the room. However, if that other person directly observes you typing, then the noise may be as tolerable to them as when producing the keyboard clicks themselves. Especially if this observer knows what you are typing or even tells you what to type, they might feel a vicarious sense of agency. Similarities and differences in the processing of outcomes of actions that are self-performed and that are observed are the topic of this section. In the following sub-sections, research on the neural activation in action observation is outlined before previous findings of auditory attenuation in action observation with perceptual and neurophysiological measures are described.

1.2.1 The Action Observation Network

The neural basis for the action observation network in the brain are mirror neurons that are active for action performance and observation and encode action goals (for reviews see Rizzolatti & Fogassi, 2014; Rizzolatti & Sinigaglia, 2010). First discovered in monkeys (e.g., di Pellegrino et al., 1992; Gallese et al., 1996), such neurons can also be found in humans (e.g., Buccino et al., 2001; Mukamel et al., 2010). In particular, a parieto-frontal network has been identified (e.g., Rizzolatti & Sinigaglia, 2010), with a meta-analysis specifying a premotor, parietal, and somatosensory network across action performance and observation, and a premotor-parietal and occipital network unique to observation (Hardwick et al., 2018). Additionally, the action observation network has been found to extend into the cerebellum (e.g., Casiraghi et al., 2019). However, the temporal sequence of activation of brain regions within the action network differed between action performance and observation: MEG studies showed, for instance, that after occipital activation the frontal and primary motor activation occurred around 100 ms later in action observation compared to action performance (Nishitani & Hari, 2000; Sebastiani et al., 2014). It was suggested that observing another's action activates mirror neurons that activate the action representation and forward models, possibly involving the cerebellum, to predict the sensory consequences (Miall, 2003). Importantly, both forward models and predictive coding proposed the same resource for predicting action outcomes in action performance and observation, for which such simulations via the action observation network might be used (Kilner et al., 2007; Pickering & Clark, 2014; Wolpert et al., 2003).

The action observation network thus supposedly enables action matching, which might be useful for action imitation, and goal matching, which can be useful for action understanding without inference, though inference is needed for understanding the reasons of the observed person (e.g., Rizzolatti & Sinigaglia, 2010; see also Urgesi et al., 2014). Such inferences are made with the so-called mentalising system, which appeared to be independent from the mirror system (Van Overwalle & Baetens, 2009). Mirror neurons have further been linked to affective understanding and empathy (e.g., Corradini & Antonietti, 2013; Iacoboni, 2009; but see Lamm & Majdandzic, 2015).

1.2.2 Perceptual Attenuation in Action Observation

As described in Section 1.1.1, paradigms for auditory perceptual attenuation typically include a tone-comparison task. In the study by Sato (2008), participants either generated the first tone themselves or observed the experimenter generating it. After the second tone (i.e., the comparison tone) was presented, participants judged which tone was louder and based on their responses the PSE was determined. Their first experiment implemented unpredictable external tones as the comparison condition. After finding significant attenuation in action performance and observation, they added a comparison condition, in which a robot arm generated the first tone, in their second experiment. They found significant attenuation in action performance and observation to both comparison conditions and concluded that the action-observation attenuation was not merely due to predictability based on visual cues. Explicit measures of agency were also obtained and found to be self-specific, in that no agency was felt for tones generated by the experimenter (Sato, 2008). To refine the approach of this study, Weiss et al. (2011a) conducted two experiments with learning phases for both action performance and observation: The first experiment applied distinct tone frequencies for action performance and observation, whereas the second applied the same tone frequency throughout. A 50 ms delay was implemented between button press and tone onset. Furthermore, the comparison condition was altered to a visual cueing condition, in which a cue was presented for 500 ms with a 100 ms interval to tone onset. In contrast to Sato (2008), Weiss et al. (2011a) consistently found attenuation only in action performance. This was the case even when the observed action was cued and therefore temporally predictable (Weiss & Schutz-Bosbach, 2012).

As a potential explanation, Cao and Gross (2015) then explored the discrepant findings in action observation as an effect of culture. Specifically, they explored it as an effect of the individualism-collectivism dimension, because attenuation in action observation was found for the Japanese (Sato, 2008), a (moderately) collectivistic culture (Hofstede, 2011), but not for the German sample (e.g., Weiss et al., 2011a), an individualistic culture (Hofstede, 2011). Individualism-collectivism describes “the degree to which people in a society are integrated into groups” (Hofstede, 2011, p. 11). At the individual level, individualism and collectivism are considered as two separate scores (Hofstede, 2011), which do not necessarily differ (both high: bi-cultural, both low: a-cultural; Singelis, 1994; Yamada & Singelis, 1999). Cao and Gross (2015) acquired a British and a Chinese sample and collected scores of the self-construal scale (SCS; Singelis, 1994), that is, independent (i.e., individualistic) and interdependent (i.e., collectivistic) SCS scores, from the participants. Their findings supported their hypothesis of an effect of culture, in that Chinese participants showed attenuation in action observation, but British participants did not, and attenuation values (i.e., PSE difference values) correlated with (independent) SCS scores (Cao & Gross, 2015). Thus, the British, and possibly the Germans in Weiss et al. (2011a), were too individualistic to attenuate the loudness of tones generated by another person they observed.

1.2.3 Neurophysiological Attenuation in Action Observation

In an EEG study, Poonian et al. (2015) investigated auditory attenuation in action observation as well as another phenomenon termed an implicit measure of agency. Showing videos of keypresses led to N1 attenuation (i.e., compared to a control condition) in action observation similar to action performance (Poonian et al., 2015). In a second experiment, they further showed reduced N1 amplitudes in action observation to observation of an unrelated action. However, the interval-recreation task and the variable delay of 500 to 1500 ms between (observed) action and tone reduce comparability to studies applying the contingent paradigm with tones following actions immediately or with a fixed delay.

Acquiring a sample of performers and a sample of observers with the contingent paradigm, Ghio et al. (2018) also found neurophysiological attenuation in action observation. This paradigm contained additional external tones that were presented after the generated tone in 40% of the trials at various delays (i.e., intermixed external tones). Notably, adding the duration of the onscreen button-press movement before the button press and the short delay between button press and tone, a participant observing the video could predict that the tone would follow 175 ms after movement onset. N1 and P2 amplitudes of tones generated by an observed action were attenuated relative to external tones, but differences in attenuation to action performance were noted. Firstly, whereas N1 attenuation in action performance was found relative to intermixed external tones, N1 attenuation in action observation was not significant in this comparison. Secondly, P2 attenuation in action performance was significantly stronger than the P2 attenuation in action observation. These findings were extended by Ghio et al. (2021) who acquired performers and observers simultaneously (i.e., the observers observed the performers live). There was again a short delay between button press and tone but a participant observing the participant next to them faced variable (and possibly shorter) movement durations. On the one hand, the N1 was attenuated only in action performance, not in action observation, which considering the results of the previous study was related to the lack of temporal predictability (Ghio et al., 2021; Ghio et al., 2018). On the other hand, the P2 was attenuated in action performance and observation and did not differ significantly between the two (Ghio et al., 2021). The N1 and P2 were interpreted as contributing to the feeling of agency and to agency judgement, respectively (Ghio et al., 2021; Ghio et al., 2018; Synofzik et al., 2008).

One limitation of these studies was that they did not control for temporal predictability, which was identified as a confound in attenuation studies (Hughes et al., 2013) and was shown to affect attenuation (e.g., Kaiser & Schütz-Bosbach, 2018), though this is not consistently shown (e.g., Klaffehn et al., 2019). Using a different approach, van Laarhoven et al. (2021) investigated the effect of temporal (and identity) predictability on N1 and P2 amplitudes, when observing handclaps onscreen. Clap sounds were presented at the end of a 500 ms handclap motion, shifted by variable delays between 250 and 170 ms before, and between 210 and 320 ms after, or substituted by random sounds. N1 and P2 attenuation was found in all conditions. For the N1, attenuation relative to external tones was stronger with clap sounds presented on time (i.e., when onset and identity were predictable). For the P2, attenuation relative to external tones was selectively decreased when the timing was unpredictable. Thus, temporal predictability was relevant for both N1 and P2 attenuation in action observation and should be accounted for in future studies.

1.3 Aims and Research Questions: Action Performance vs. Observation

This dissertation investigated auditory action-outcome processing of action performance and action observation in three studies. Specifically, the attenuation of N1 and P2 ERP components of action-generated tones was compared to that of externally generated tones. In previous studies, the condition to which attenuation in action observation was examined either was the typical auditory-only condition in neurophysiological attenuation studies (i.e., unpredictable external tones; e.g., Ghio et al., 2021; Ghio et al., 2018) or included visual cues. For instance, a visual cue preceded the externally generated tone by 600 ms in Weiss et al. (2011a). Several of the described studies presented the tones immediately after action-onset (e.g., Cao & Gross, 2015; Sato, 2008) or the delay between action and outcome was very short (i.e., 50 ms; e.g., Ghio et al., 2021; Weiss et al., 2011a). In Ghio et al. (2018) the delay between movement onset of the observed button press and the tone was longer (i.e., 175 ms), supposedly enabling better temporal predictability for observed actions and thus the generated tone (see Ghio et al., 2021; Ghio et al., 2018). However, studies implementing live action observation would also feature time delays between movement onset and full button press and thus the tone, which likely vary inter-individually and between trials.

One overarching aim of the dissertation, was to examine whether the previous attenuation effects in action observation (Ghio et al., 2021; Ghio et al., 2018) were due to temporal predictability. To account for the effect of temporal predictability, we implemented (i) a larger action-outcome delay making the tones of observed actions temporally predictable, and (ii) a condition with external tones that were made temporally predictable by visual cues in all three studies. Each of the three studies then addressed a specific research question.

The specific research question of Study 1 was to investigate the potential effect of culture (i.e., individualism) on neurophysiological attenuation in action observation as measured by the N1 and P2 amplitudes. In the study by Cao and Gross (2015), the participant's culture determined the degree of perceptual attenuation in action observation. Traits were measured by questionnaires and correlated with attenuation values in action observation. A decreasing effect on PSE attenuation in action observation was found for the independent SCS score (i.e., individualism), while no (increasing) effect of the interdependent SCS score (i.e., collectivism) was found (Cao & Gross, 2015). Individualism has not yet been considered as a modulatory variable in studies on neurophysiological attenuation in action observation (Ghio et al., 2021; Ghio et al., 2018). Therefore, we included an individualism-collectivism questionnaire in Study 1 to relate the scores to N1 and P2 amplitudes of self-generated tones and tones generated by an observed action in a simultaneous EEG acquisition.

The specific research question of Study 2 was to examine the role of agency attribution for the P2 attenuation in action observation. In Study 1 (Egan et al., 2023), the P2 was attenuated significantly in action performance and observation compared to the cue condition. One possible explanation for the P2 results was the relation of P2 attenuation to the sense of agency (e.g., Timm et al., 2016). In previous action-observation studies, the P2 attenuation was also interpreted in terms of agency judgements, attributing agency to oneself or to the other (Ghio et al., 2021; Ghio et al., 2018). However, non-human agents that perform the button press and generate the tone have not been included in these studies. To assess the role of agency attribution for the P2, we included robotic action-observation conditions and explicit agency measures in Study 2.

The specific research question of Study 3 entailed the effect of varying whether-prediction errors on N1 and P2 amplitudes in action performance and observation. According to the forward models, the contingent relation between action and outcome, that is, the consistent coupling of action and outcome, is a prerequisite for attenuation (Horváth, 2015). Only when the action is (causally) linked to the outcome, can predictions be made regarding its temporal onset and identity. Previous studies have reduced the contingency in one of two ways: Introducing deviants (i.e., infrequent tones with an altered frequency) reduces the contingency of action and standard tone (e.g., Knolle et al., 2013b) thereby affecting the what-prediction. Other studies have reduced the action-outcome contingency by omitting tones (for a review, see Korka et al., 2022), which affects the whether-prediction (i.e., predictions about whether the tone occurs or not). In contrast to forward-model assumptions, N1 and P2 attenuation has been found as an effect of mere temporal proximity between action and tone in studies applying the coincidence paradigm (e.g., Horváth et al., 2012), as well as N1 attenuation when there was only a 50% contingency (through omissions) between action and tone (e.g., Seidel et al., 2021; Timm et al., 2013). Focusing on whether-predictions but applying different paradigms, Han et al. (2022) and Harrison et al. (2023) compared full- (i.e., 100%) and no-contingency (i.e., 50%) conditions and found no effect on N1 amplitudes. Han et al. (2022), however, found contingency-dependent P2 reduction. Yet, it can be assumed that with uncertain contingencies the prediction varies from trial to trial and together with the (lack of) outcome results in varying prediction errors. To explore this further, in Study 3, we presented tones in only half of the trials, calculated prediction errors in each trial, and examined their effect on N1 and P2 amplitudes. Effects of contingency have not yet been investigated for action observation. Hence, in Study 3, we recruited a group who performed the button press themselves as well as a group who observed the button press onscreen.

2 Study Summaries

The three studies that were conducted are summarised in this section. For the complete articles see Appendix A (i.e., Section 6). Information on the supplementary materials are provided in Appendix B (i.e., Section 7). As stated in the introduction, the three studies investigated sensory attenuation in action performance and observation. Study 1 (Egan et al., 2023) investigated the effect of culture on attenuation in action observation based on the findings of a perceptual study (Cao & Gross, 2015). It also took the effect of temporal predictability into account that previous studies on sensory attenuation in action performance had considered (e.g., Klaffehn et al., 2019). Study 2 (Egan et al., 2025) investigated the effect of agency attribution on attenuation in action observation, which was a potential explanation of the results of Study 1 (Egan et al., 2023). Finally, Study 3 (Egan et al., 2024) investigated the effect of the prediction error on attenuation in action performance and observation by reducing the contingency between action and outcome. At the beginning of each sub-section a summary figure is provided for the study, including depictions of the action-observation setting and schematic depictions of the attenuation results.

Each study was approved by the ethics committee of the Faculty of Mathematics and Natural Sciences at the Heinrich Heine University. Participants were recruited at the university and were mostly (psychology) students. Each participant gave written informed consent and received either monetary compensation or course credit for participating. Those who reported a history of psychiatric or neurological disorders were excluded from data analysis. All studies included EEG acquisitions with 28 passive Ag/AgCl electrodes (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), AFz as ground electrode, two electrodes on the mastoids as reference, and four electrodes to record eye-movement-related and muscle activity (F9, F10, FP2, and one attached under the right eye). BrainVision Recorder was used for recording (1000 Hz sampling rate), BrainVision Analyser for pre-processing the EEG data (Brain Products GmbH). We further analysed and visualised the data with R (R Core Team, 2024). Statistical analysis in each study entailed the computation of linear mixed-effects models. These models have several advantages over a conventional analysis of variance (ANOVA): Random effects can be included in mixed-effects models (i.e., variability, such as between participants, can be modelled), they are more robust regarding missing data, and continuous variables can easily be modelled as additional predictors (Field et al., 2012). The latter advantage was exploited for adding trait scores as potential modulators in Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025) and analysing the effect of the trial-by-trial varying prediction error in Study 3 (Egan et al., 2024).

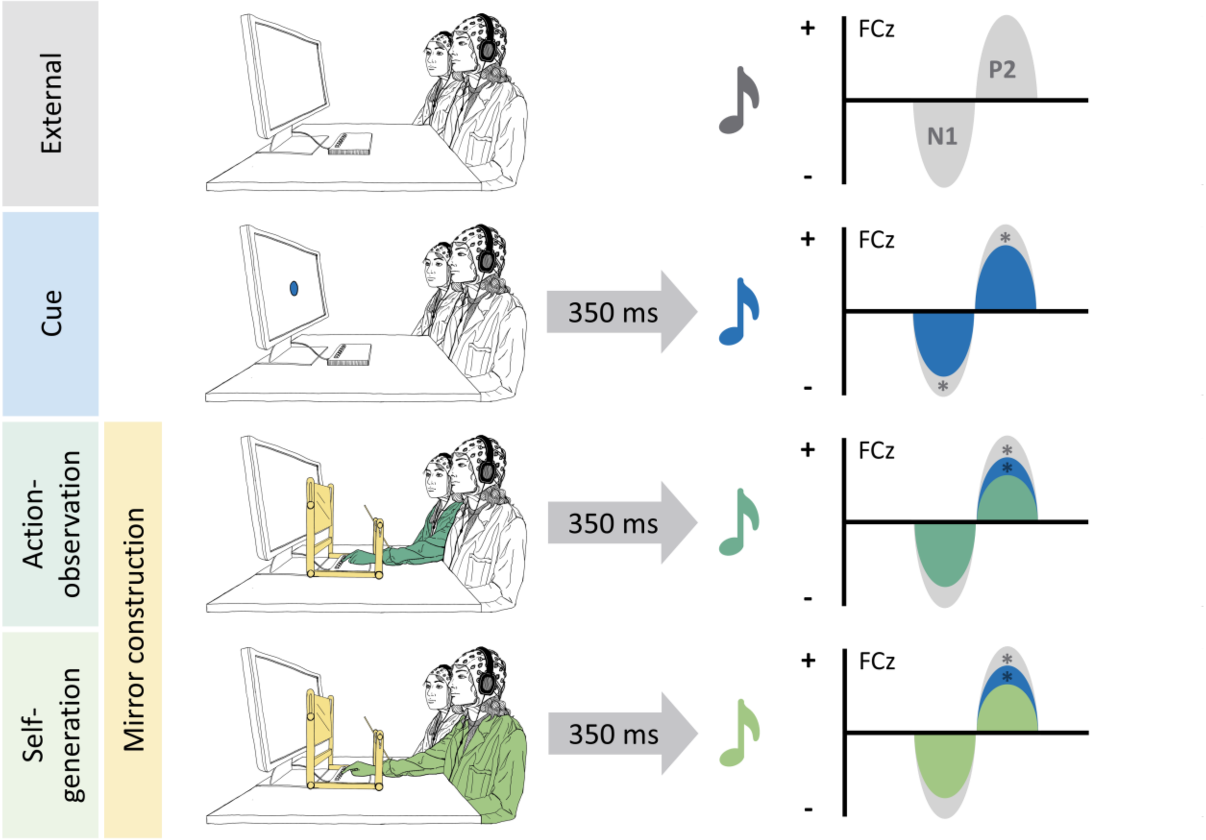
2.1 Auditory N1 and P2 Attenuation in Action Observation: An Event-Related Potential Study Considering Effects of Temporal Predictability and Individualism

The first study has been published in the journal Biological Psychology with the following reference:

Egan, S., Ghio, M., & Bellebaum, C. (2023). Auditory N1 and P2 attenuation in action observation: An event-related potential study considering effects of temporal predictability and individualism. *Biological Psychology*, *180*, 108575. https://doi.org/10.1016/j.biopsycho.2023.108575

See Figure 2.1 for an overview of the action-observation setting applied in this study and the attenuation results.

FIGURE 2.1: ACTION-OBSERVATION SETTING AND ATTENUATION RESULTS OF STUDY 1



Note. Acquisitions were done in pairs. Each participant completed four conditions, in which tones were presented. External tones were temporally unpredictable in the external-tone condition (grey) and temporally predictable in the cue condition (blue). In the action-observation condition (teal) and the action-performance condition (green), a mirror construction was placed on top of the response box, so that the button presses were viewed in first-person perspective in the mirror. To press the button comfortably with their right and left hand, participants switched seats after half of the experiment was completed. Tones followed 350 ms after button press or cue onset. For precise descriptions of the conditions see Section 6.1. On the right, N1 and P2 amplitudes from the chosen electrode are schematically depicted and significant comparisons, in particular to the external-tone condition (i.e., signifying attenuation), are marked.

2.1.1 Research Question

As mentioned in Section 1.1.2, temporal predictability has been a confound in the comparison of self-generated and externally generated tones (Hughes et al., 2013). Likewise, it presents a confound for attenuation in action observation, as the observed action might merely serve as a cue to make the tones predictable in time. The effect of temporal predictability has previously been investigated by making externally generated tones temporally predictable using visual cues, for example, to compare them to self-generated tones (e.g., Klaffehn et al., 2019). Results indicated that while N1 amplitudes of self-generated tones can be attenuated compared to temporally predictable external tones (e.g., Klaffehn et al., 2019), N1 and P2 amplitudes of temporally predictable external tones can also be attenuated compared to unpredictable external tones (e.g., Schafer & Marcus, 1973).

The findings of perceptual attenuation in action observation described in Section 1.2.2 were inconsistent and a cultural dimension, namely individualism, was suggested as a possible explanation for the discrepancies (Cao & Gross, 2015). Specifically, attenuation in action observation was found for a Japanese (Sato, 2008) and a Chinese sample (Cao & Gross, 2015), both cultures considered to be more collectivistic, but no attenuation was found in observers for a German (Weiss et al., 2011a) and a British sample (Cao & Gross, 2015), both cultures considered to be more individualistic. In the study by Cao and Gross (2015), scores measuring individualism did indeed correlate with attenuation values (i.e., differences of PSE values in observation and external conditions). While neurophysiological measures were also (partially) attenuated in action observation for German samples (Ghio et al., 2021; Ghio et al., 2018), the modulation by individualism has not yet been examined. Importantly, the external-tone condition in these perceptual studies included a visual cue that made the external tones temporally predictable (e.g., Weiss et al., 2011a), but the external tones in the neurophysiological studies were not (Ghio et al., 2021; Ghio et al., 2018).

The first study of this dissertation therefore investigated the attenuation of N1 and P2 amplitudes in action observation by expanding the previous paradigm of Ghio et al. (2021) with a control condition for temporal predictability (i.e., a condition with cued external tones) matched in timing to the conditions of action performance and observation. In this way, the question of the relevance of temporal predictability in attenuation in action observation was addressed (Ghio et al., 2021; Ghio et al., 2018). Regarding the action-observation condition, participants were acquired simultaneously as in Ghio et al. (2021), but each participant performed and observed button presses. To address the question of whether individualism affects the attenuation in action observation as suggested by perceptual studies (Cao & Gross, 2015), participants filled out the Auckland individualism collectivism scale (AICS; Shulruf et al., 2011). The AICS measures individualism and collectivism as distinct dimensions, providing a separate score for each. We then tested whether the addition of the respective score significantly improved the models predicting N1 and P2 amplitudes. There were thus two research aims: The first was to compare action observation to cueing through other visual stimuli and the second was to explore the effect of culture. We expected to find attenuation of N1 and P2 amplitudes in action performance and observation (Ghio et al., 2018; Wolpert et al., 2003) compared to cued and un-cued external tones (e.g., Klaffehn et al., 2019, for action performance), assuming action-related mechanisms beyond temporal predictability contribute to attenuation. Based on the perceptual study by Cao and

Gross (2015), we further expected an interaction with individualism, as the expected attenuation in action observation might be particular to participants with low individualism scores.

2.1.2 Methods

Our experiment had one within-subject factor encompassing four tone-generation conditions and the action-observation setting in this study was live, in that two participants were acquired simultaneously (see also Ghio et al., 2021). Both participants performed button presses themselves and each observed the other pressing the button, which is a difference to previous studies acquiring a group of performers and a group of observers (Ghio et al., 2021; Ghio et al., 2018). To increase comparability between performance and observation, a mirror construction reflected the button press to the height of the screen in a first-person perspective (for details see Figure 2.1). In a learning phase, both participants took turns in two tasks to learn the rhythm, in which they were to press the button, and to familiarise themselves with the action-effect, namely, the tone. They pressed the button with their right and left index finger in respective blocks, but we did not differentiate these in the action-performance and action-observation conditions. Together with the cue and un-cued external-tone conditions, the experiment had four conditions (see Figure 2.1). A delay of 350 ms between event onset (i.e., button press or cue presentation onset) ensured temporal predictability in all but the un-cued external-tone condition. Generated tone sequences from the action-performance and -observation conditions were played back in the cue and external-tone conditions. Likewise, cue presentation duration was matched to the button press duration in those conditions. As in the contingent paradigm, control blocks were included to later correct the ERPs of tone-unrelated activity (cf. motor correction; Horváth, 2015). Thus, there were motor-only blocks for action performance and observation and a visual-only block for the cue condition (i.e., presenting the cues without tones). Each block comprised 120 trials and was divided into two sub-blocks (which were completed with right and left index finger, respectively, in action performance and observation blocks). The order of blocks could not be fully randomised because of the timing dependency (i.e., of intervals and button press durations) and the setup with the mirror construction (see Figure 2.1). Instead, four versions of the experiment with counterbalanced orders were created.

The potentially specific effect of culture on attenuation in action observation was assessed via a questionnaire, the AICS (Shulruf et al., 2011), to be used as a predictor for N1 and P2 amplitude values. The questionnaire yielded an individualism score with the sub-scores competitiveness, uniqueness, and responsibility; and a collectivism score with the sub-scores advice and harmony. The aim was to have variance in individualism scores, for which participants from different nationalities were recruited. As empathy and a concept similar to anomalous perception were also considered in the perceptual study (Cao & Gross, 2015), anomalous perception (e.g., perceiving sounds as louder than they would normally be) measured by the Cardiff anomalous perceptions scale (CAPS; Bell et al., 2006) and empathy measured by the empathy quotient (EQ; Lawrence et al., 2004) were also considered as potential modulators of N1 and P2 attenuation.

After excluding eight participants due to technical issues and a minimal trial-number criterium (i.e., 60% usable trials in every block), the sample comprised 49 participants. The majority was of German nationality (39, including three dual nationalities). The remaining participants were of other European or Asian nationality. Trials were excluded during the pre-processing of the EEG data, which included applications of direct current detrend, filters, ocular correction independent component analysis, segmentation, baseline correction, and artifact rejection (see Section 6.1 for details). Participant averages were calculated and corrected (i.e., the averages of control motor/visual-only blocks were subtracted from the averages of auditory-motor/visual blocks, respectively). The FCz electrode showed the highest N1 and P2 peaks (in line with previous studies, e.g., Ghio et al., 2021) and was thus chosen for further analysis. FCz amplitude values were averaged in the interval from 92 to 112 ms for the N1 and in the interval from 161 to 211 ms for the P2 for each participant and condition. The intervals were determined based on the latencies of the grand average (i.e., collapsed localiser technique; Luck & Gaspelin, 2017) and interval length was based on visual inspection of the widths of the peak regions of the condition averages.

The basic linear mixed-effects model for N1 and P2 included the fixed effect of condition and a random intercept by participant: amplitude ~ condition + (1 | participant). Outliers were determined with Cook's distance. This resulted in the exclusion of two additional participants for the P2 analysis only (i.e., 47 participants included). An ANOVA-like analysis was applied on the mixed-effects models and post-hoc pairwise comparisons were Bonferroni-corrected. Regarding the second research aim, the exploration of the effect of culture, the models for N1 and P2 were extended and then the models with and without the questionnaire score were compared to determine if the addition of the questionnaire score significantly improved the model: amplitude ~ condition * questionnaire score + (1 | participant). This was done for the AICS individualism and the AICS collectivism score as well as for each sub-score, as these allow for finer cultural differentiation (Shulruf et al., 2007). Additionally, we calculated and compared the models for the EQ and CAPS scores (cf. Cao & Gross, 2015).

2.1.3 Results and Discussion

A schematic summary of the results can be found in Figure 2.1. The effect of condition was significant for the models of N1 and P2 amplitudes. Only the N1 amplitudes of cued external tones were significantly attenuated compared to the un-cued external tones, while the N1 amplitudes were descriptively reduced in action performance and observation. All other pairwise comparisons were not significant. The P2 amplitudes of all temporally predictable tones were significantly attenuated compared to the un-cued external tones. P2 attenuation in action performance and observation did not differ significantly and P2 amplitudes for both conditions were significantly attenuated compared to cued external tones. The addition of the AICS individualism score did not improve model fit for N1 or P2. Other questionnaire scores, such as AICS collectivism, the respective AICS sub-scores, the scores derived for anomalous perception (i.e., CAPS scores), and the EQ score did not improve model fit for N1 or P2 either. Therefore, the extended models were not further analysed.

An explanation for the results of the N1 amplitude might be temporal predictability, particularly because cued external tones were significantly attenuated and did not differ significantly from amplitudes in action performance and observation. This aligns with some previous studies on attenuation in action performance (e.g., Harrison et al., 2021) while contrasting with others (e.g., Klaffehn et al., 2019). Alternatively, the tone-preceding events in the conditions with temporal predictability could have diverted attention from the tones and reduced the N1 amplitudes (Horváth, 2015) but previous studies testing this attention hypothesis found no support for it (Harrison et al., 2021;

Timm et al., 2013). Surprisingly, N1 amplitudes of self-generated tones were not significantly attenuated, despite it being a common finding (Horváth, 2015), even with action-outcome delays (e.g., Lange, 2011). Future studies might investigate whether the altered view of one's own hand, which the mirror construction in this study provided, affects N1 attenuation.

The P2 attenuation showed an effect of temporal predictability (i.e., attenuation for temporally predictable tones) and of action (i.e., stronger attenuation when action-generated) with no difference between performance and observation. However, it is possible that the (self-performed or observed) action simply was a better predictor than the visual cues (i.e., the circle that appeared 350 ms before the tone), considering that (action planning and) finger movement began earlier than the presentation of the cues. Temporal predictability was shown to affect P2 attenuation in a previous study with clap sounds (van Laarhoven et al., 2021) but other studies did not find P2 attenuation for delayed tones (Klaffehn et al., 2019) or cued external tones (Harrison et al., 2021). On the other hand, several studies related the P2 attenuation to the sense of agency, particularly the agency judgement (e.g., Timm et al., 2016). Accordingly, our P2 results might suggest that the generation of the tones was attributed to oneself or the other person, in other words, a human agent, but was not attributed or less so to a computer-generated cue. Whether this is indeed specific to human agency or due to other distinctions to the visual cueing in our study could be further examined and was subsequently addressed in Study 2.

Unlike the study by Cao and Gross (2015) on perceptual attenuation, we did not find an effect of individualism on N1 or P2 amplitudes, as individualism did not significantly improve the models. In previous perceptual studies, attenuation in action observation was not found consistently (Sato, 2008; Weiss et al., 2011a) and a cultural effect, namely the degree of individualism, seemed to explain this inconsistency, in that more individualistic participants as observers showed less attenuation (Cao & Gross, 2015). However, few studies have compared perceptual and neurophysiological measures of attenuation and questioned the relationship between the two measure types (e.g., Palmer et al., 2016). Furthermore, whereas Cao and Gross (2015) recruited two groups of participants with different nationalities (one considered individualistic, and one considered collectivistic), our sample mainly consisted of participants scoring similarly on individualism and collectivism, though scoring across the range of the scores, rather than scoring high on one and low on the other.

Aside from the sample possibly being too moderate to discover the effects of the acquired questionnaire-based measures, another limitation of our study is the applied motor correction (see Horváth, 2015). This correction could be inadequate and affect the observed attenuation for N1 and P2 (Neszmélyi et al., 2022).

In sum, we extended the findings of sensory attenuation in action observation. Participants observed the hand of the other participant, who was acquired simultaneously, in first-person perspective via a mirror construction. We did not find any differences in N1 and P2 amplitude between action performance and action observation. Temporal predictability played a role for N1 and P2 attenuation as evident by the attenuation for cued external tones compared to un-cued external tones. For the P2, however, it could not explain the further attenuation in action performance and observation, which might relate to the sense of agency, experienced for oneself or vicariously. Potential modulations on N1 or P2 attenuation by individualism, collectivism, anomalous perception, and empathy considered via questionnaire scores could not be supported in this study. Regarding the two research aims, it can be said that the attenuation in action observation went beyond an effect of temporal predictability for the P2 and that unlike the perceptual measure of perceived loudness (i.e., PSE), the N1 and P2 seemed unaffected by individual differences on the measured cultural dimension.

2.2 I, You, Robot: Attenuation for Auditory Outcomes of Actions Performed by Different Agents Shows Distinct Patterns for N1 and P2 Amplitudes

The second study has been submitted for publication:

Egan, S., Weber, C., Ghio, M., & Bellebaum, C. (2025). I, You, Robot: Attenuation for Auditory Outcomes of Actions Performed by Different Agents Shows Distinct Patterns for N1 and P2 Amplitudes [Manuscript submitted for publication].

See Figure 2.2 for an overview of the action-observation setting applied in this study and the attenuation results.

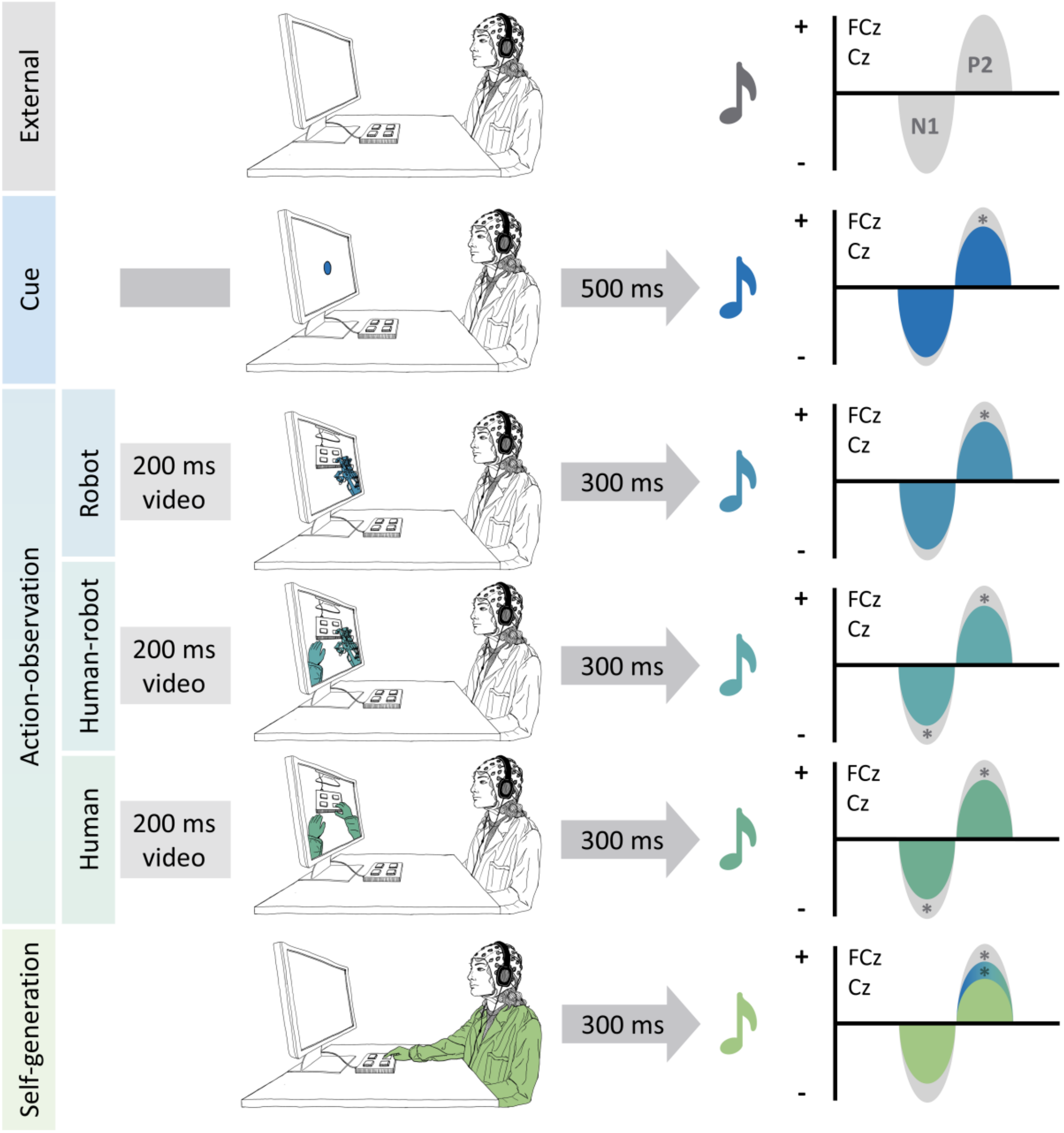
2.2.1 Research Question

In Study 1 (Egan et al., 2023), we found attenuated P2 amplitudes in action performance and observation compared to cued external tones. As has been described in Section 1.1.3, P2 attenuation was previously related to agency judgements (Timm et al., 2016), and P2 amplitudes were reduced when participants had a high rather than low illusion of control over the tone (Seidel et al., 2021). As in action performance, an agent is involved also in action observation. The agency for tone generation may thus be attributed to oneself or to the observed person (Ghio et al., 2021; Ghio et al., 2018). This would be in line with the similar P2 attenuation in action performance and observation of Study 1 (Egan et al., 2023). However, it could alternatively be an effect of the learnt action-outcome association that was missing in the cue condition, which instead presented an association of an external event (i.e., appearance of a circle) and the outcome (i.e., tone).

A comparison condition with an action by a non-human agent, such as a robot, could provide the same action-outcome association as in action observation but without a human agent. Interestingly, such actions can also activate the mirror neuron system (Van Overwalle & Baetens, 2009). However, in the perceptual study that found attenuation in action observation and included a comparison condition with a robotic arm generating the tone, this was not supported (Sato, 2008). Specifically, the PSE did not differ from the PSE of the comparison condition with external tones, while the PSE values in action performance and observation were significantly attenuated also relative to the robotic condition. This was interpreted as highlighting the relevance of a human agent for (activation of the mirror neuron system and) attenuation (Sato, 2008).

In Study 2, we thus included action-observation conditions that differed in attribution to a human or robotic agent, to investigate the role of agency attribution for P2 attenuation in action observation. Specifically, videos of button presses performed by a human hand, a human-operated robotic hand, and a computer-controlled robotic hand were shown.

FIGURE 2.2: ACTION-OBSERVATION SETTING AND ATTENUATION RESULTS OF STUDY 2



Note. Each participant completed six conditions, in which tones were presented. External tones were temporally unpredictable in the external-tone condition (grey) and temporally predictable in the cue condition (blue). In the action-observation conditions, participants observed the button press onscreen in first-person perspective performed by a robotic hand (muted blue), a robotic hand with human agent (cool teal), and a human hand (teal). In the action-performance condition (green), participants pressed the button themselves. Tones followed 300 ms after button press and 500 ms after cue onset. Note that in action-observation conditions there was an additional interval from movement onset till button press. For precise descriptions of the conditions see Section 6.2. On the right, N1 and P2 amplitudes from the chosen electrodes are schematically depicted and significant comparisons, in particular to the external-tone condition (i.e., signifying attenuation), are marked.

Given the effects of temporal predictability on N1 and P2, the comparison condition with cued external tones from Study 1 (Egan et al., 2023) was retained and matched in timing to the action-observation conditions. Thus, the primary research aim was to compare attenuation for action observation of human and of robotic agents, as this could offer further insights on the results of Study 1 (Egan et al., 2023). The secondary research aims were to explore the effects of various participant traits on attenuation and thus account for potential biases. We expected to find N1 attenuation for all temporally predictable tones relative to unpredictable external tones, in line with previous studies (e.g., Harrison et al., 2021), although this was found only for cued external tones in Study 1 (Egan et al., 2023). For the P2, we expected attenuation for tones attributed to a human agent relative to those attributed to the computer but also an effect of temporal predictability (Egan et al., 2023). That is, if the P2 attenuation in action observation compared to cued external tones found in Study 1 (Egan et al., 2023) relates to the human agency, we should replicate this finding for the human and human-robot action-observation compared to robot action-observation and cue conditions. However, for participants with a high tendency to anthropomorphise (i.e., attributing human characteristics to non-human animals and objects), the robot action-observation condition might be more akin to the human and human-robot action observation conditions and attenuated compared to the cue condition. Therefore, we included scores of the anthropomorphism questionnaire (AQ; Neave et al., 2015) and expected an interaction.

2.2.2 Methods

As in Study 1 (Egan et al., 2023), we had one within-subjects factor encompassing the six different tone-generation conditions, but the action-observation setting was altered. In the action-observation conditions, videos of the button press in first-person perspective were presented onscreen (see Figure 2.2), so that timings could be better controlled. The action-observation conditions included (i) one with a human agent and natural depiction (i.e., human hand), (ii) one with a human agent and robotic depiction (i.e., robotic hand), and (iii) one with a non-human agent with robotic depiction. Thus, human-robot and robot action-observation conditions differed solely in agency attribution, and human and human-robot action-observation conditions differed solely in visual likeness to one's own hand. At the end of the experiment, participants evaluated the depiction on scales regarding the overall impression (e.g., fake vs. natural) and the movement (i.e., rigid vs. elegant). Similarly to Study 1 (Egan et al., 2023), we set a delay of 300 ms between the picture showing the button completely pressed and the tone to ensure temporal predictability (350 ms in Study 1). To obtain comparable inter-tone intervals to self-generation, the intervals between button-press videos were determined by a function that randomly chose an interval from a normal distribution (*M* and *SD* based on previous data). However, there was a latency error in stimuli presentation, and each picture was presented longer than intended, which prolonged the calculated intervals. Likewise, the button press video was prolonged and had a duration of 200 ms. The cue condition was matched in timing to the movement onset and offset in the action-observation conditions (i.e., 500 ms between cue onset and tone). Participants also pressed the button themselves in an action-performance condition and listened to a tone sequence of unpredictable external tones in the external-tone condition. Thus, we had six conditions in total (cf. Figure 2.2). In the action-performance condition, participants self-timed the button presses according to the rhythm they learned in a training at the beginning of the experiment and the tone was presented 300 ms after the button press.

In the external-tone condition, the inter-tone intervals were generated randomly from a normal distribution approximating the participant's self-timed intervals. As in Study 1 (Egan et al., 2023), we included control blocks without tone presentations for action-performance, action-observation, and cue conditions to later correct the ERPs of tone-unrelated activity (cf. motor correction; Horváth, 2015). Each block had 80 trials. Blocks were grouped in experimental and control blocks. Participants began with either the first group of blocks or the latter, and block order within the group was randomised, respectively.

Four questionnaires assessed participant traits and were tested for a modulation of attenuation. As in Study 1 (Egan et al., 2023), this included anomalous perception measured by the CAPS (Bell et al., 2006) and empathy measured by the EQ (Wakabayashi et al., 2006). The use of the robotic depiction in two action-observation conditions led us to further include anthropomorphism measured by the AQ (Neave et al., 2015) and the general attitudes towards robots scale (GAToRS; Koverola et al., 2022). Additionally, because of the relevance of the sense of agency to our primary research aim, we collected an explicit measure of agency after each experimental block: a judgement (i.e., who caused the tone) and an agency-confidence rating (i.e., how strongly did they feel their proclaimed agent was indeed the agent of the tone).

The final sample comprised 40 participants, after nine exclusions due to lacking data for one experimental block (i.e., missing or excluded during pre-processing). Pre-processing of the EEG data followed these steps: direct current detrend, filters, ocular correction independent component analysis, manual marking of bad intervals due to closed eyes, segmentation, baseline correction, and artifact rejection (see Section 6.2 for details). Trials were excluded at segmentation and artifact rejection. Further, an entire experimental block and thus the participant was excluded, when the participant performed button presses in the action-observation, cue, or external-tone conditions. For the correction of tone-unrelated activity, the single-trial ERPs from the experimental blocks of action-performance, action-observation, and cue conditions were corrected with the averages of the respective control blocks. The electrodes FCz and Cz showed the most pronounced N1 and P2 peaks in grand averages (cf. Klaffehn et al., 2019) and were thus chosen for further analysis. Peak latencies were determined based on condition-specific grand averages. For action-performance, human, human-robot, robot, cue, and external-tone conditions, amplitudes were averaged in 20-ms intervals around the peaks for the N1 on FCz (99, 87, 87, 83, 84, and 91 ms) and Cz (98, 86, 86, 82, 83, and 90 ms), and in 40-ms intervals around the peaks for the P2 on FCz (192, 158, 160, 162, 162, and 170 ms) and Cz (195, 155, 159, 161, 152, and 171 ms).

As a first step, mixed-effects models were calculated on the behavioural data regarding agency judgement and agency-confidence ratings because several participants gave incorrect agency judgements (e.g., claimed themselves to be the agent in the external-tone condition) and the agency-confidence ratings were meant to be added to N1 and P2 models. Inaccuracies were found primarily for the action-observation conditions, but the overall pattern of judgement responses was as intended (e.g., judged the computer as the agent in the external-tone condition). The model for the analysis of agency-confidence ratings included only the condition as fixed effect as the further addition of agency judgment as predictor did not improve the model significantly: rating ~ condition + (1 | participant). We then calculated Bonferroni-corrected pairwise comparisons.

The primary analysis was conducted using linear mixed-effects models of N1 and P2 amplitude values. For the N1, we included condition as the fixed effect as well as random intercept and slope by participant: amplitude ~ condition + (condition | participant). The model for the P2 further included random intercept by electrode because this significantly improved the model: amplitude ~ condition + (condition | participant) + (1 | electrode). ANOVA-like analyses and Bonferroni-corrected pairwise comparisons were applied. For the secondary analysis, the models for N1 and P2 were extended by a trait-score main effect and its interaction with condition. The models with the obtained questionnaire scores (four from CAPS, one EQ, three from AQ, four from GAToRS) were compared to the model without. Finally, the models for N1 and P2 were similarly extended sequentially by agency-confidence ratings and judgement accuracy, and the models were compared.

2.2.3 Results and Discussion

The analysis of agent-confidence ratings confirmed the results from the analysis of the agency judgement inaccuracies, in that confidence ratings were lower in action-observation than in the other conditions, particularly for human-robot (and robot) action observation. A schematic summary of the N1 and P2 results can be found in Figure 2.2. The effect of condition was significant in both. Compared to the external-tone condition, only the human and human-robot action-observation conditions showed significant N1 attenuation. Other pairwise comparisons were not significant. On the other hand, the P2 in action-performance, action-observation, and cue conditions was significantly attenuated compared to the P2 of temporally unpredictable external tones. The P2 in action performance was further significantly attenuated compared to the P2 in action-observation and cue conditions. There were no significant differences between any of the action-observation and cue conditions. Regarding the model comparisons, there was no trait score that significantly improved the N1 and P2 models. The subsequent addition of agency-confidence rating and agency judgement did not improve the N1 and P2 models either. Therefore, the extended models were not further analysed.

Since the results of Study 1 indicated that N1 attenuation reflects temporal predictability and that temporal predictability could further explain part of the P2 attenuation in action observation (Egan et al., 2023), we accounted for the effect of temporal predictability in Study 2 as well. Interestingly, in this study we found no differences between action observation and cueing in either the N1 or the P2. Thus, the attenuation in action observation relative to unpredictable external tones found in this study was seemingly based solely on temporal predictability. However, this attenuation of the N1 was significant only for tones generated by actions attributed to a human agent. This, in turn, raises doubts about the interpretation of N1 attenuation as reflecting only temporal predictability, as previous studies have also done (Klaffehn et al., 2019; Schafer & Marcus, 1973). For the P2 that showed an additional attenuating effect of action-performance, our results stand in contrast to our hypotheses that attenuation in action observation would be as in action performance and stronger than for cued external tones (see Egan et al., 2023). A difference to Study 1 (Egan et al., 2023) was the action-observation setting. We changed it to video observation for better control of the timing, which could have reduced the P2 attenuation. In line with this explanation, a previous study implementing video observation also found weaker P2 attenuation in action observation than performance (Ghio et al., 2018).

What might modulate the magnitude of P2 attenuation in action observation – akin to action performance in a live-observation setting and akin to cueing in a video-observation setting – is the perception of agency. The P2 attenuation has been previously related to (vicarious) agency (Ghio et al., 2021; Timm et al., 2016). Yet, effects of self-agency were shown to be stronger on the PSE in German samples (Weiss et al., 2011a, 2011b). This may in part rely on an effect of temporal control, which is specific to action performance and which was also found to be relevant to P2 attenuation: In the comparison of self-timed and cued actions, only tones of self-timed action showed P2 attenuation (Harrison et al., 2021). In our study, we acquired explicit agency measures (i.e., judgement who the agent was and rating how strongly the chosen agent was perceived as the agent). Although weaker agency attribution was found for the action-observation conditions, the addition of the explicit agency measures did not improve the P2 model. The relation between P2 and explicit agency attribution could thus not be convincingly supported.

Intentional binding (i.e., perceived timing shift of action and outcome towards each other) has also been regarded as an implicit measure of the sense of agency. Studies have shown the binding effect for action observation (e.g., Poonian & Cunnington, 2013), including action observation of robotic agents, the results of which suggested that human-likeness (in physical settings) and perceived intentionality were relevant factors in this regard (Roselli et al., 2022; Roselli et al., 2021). In our study, participants rated the action-observation stimuli regarding naturalness, human-likeness, lifelikeness, and smoothness of movement. In all ratings, the human hand scored significantly higher than the robotic hand, although in the latter rating the human hand had low ratings, too. The rating differences were not reflected in the P2 amplitudes in this study but might be reflected in P2 amplitudes in a live action-observation setting (Ghio et al., 2021; Roselli et al., 2022). Furthermore, the ratings indicating rigid movement were likely due to the presentation-latency error slowing the video of the button press and reducing the perceived similarity in motion to the participants' own actions. This in turn was proposed to lead to less anthropomorphism and less activation in the mirror neuron system (Epley et al., 2007), which could be another reason for the diminished P2 attenuation in action observation.

Surprisingly, we did not find N1 attenuation in action performance (see also Egan et al., 2023), although it is usually reported (Horváth, 2015), also in studies implementing longer action-outcome delays (e.g., Klaffehn et al., 2019). It should be noted that in the contrasts of the fixed-effect estimates with uncorrected *p*-values, there was a significant N1 attenuation in action performance. However, the confidence interval was larger in comparison to other conditions, which could be due to the additional attentional demands of the task. Another potential confound of the N1 was related to the differences in inter-tone intervals (SanMiguel, Todd, et al., 2013). These were longer particularly in the action-observation and cue conditions relative to the external-tone condition because of the presentation-latency error, thus possibly reducing attenuation, whereas self-generated tones had a mean value between those two. When considering the participant means of inter-tone-intervals, self-generated tones further showed a larger standard deviation, which could be another explanation for the non-significant attenuation.

Furthermore, we did not find an improvement of N1 or P2 models when adding trait scores (anomalous perception, empathy, anthropomorphism, and attitudes towards robots) as predictors. This is in line with the results of Study 1 (Egan et al., 2023), and regarding empathy in line with findings on PSE attenuation, though PSE attenuation did correlate with the PDI scores for delusional ideation (Cao & Gross, 2015). Thus, the considered traits did not seem to influence N1 and P2 attenuation.

In sum, we investigated whether the P2 attenuation of Study 1 (Egan et al., 2023) could be explained by agency attribution. For this purpose, participants watched videos of button presses performed by a human hand or a robotic hand attributed to a human or non-human agent. The N1 was significantly attenuated only in action observation with human agents but did not differ between action performance and all other conditions. P2 attenuation compared to unpredictable external tones was found for temporally predictable external tones regardless of action-observation or agency and for self-generated tones, which was significantly stronger than for the predictable external tones. Regarding the first research aim, therefore, we did not find an effect of agency attribution on the P2 attenuation. However, the action-observation setting was different compared to Study 1 (Egan et al., 2023). Concerning the secondary research aims, we found, similarly to Study 1 (Egan et al., 2023), no significant effects of traits on N1 or P2 attenuation.

2.3 Fifty Percent of the Time, Tones Come Every Time: Stronger Prediction Error Effects on Neurophysiological Sensory Attenuation for Self-Generated Tones

The third study has been published in the Journal of Cognitive Neuroscience with the following reference:

Egan, S., Seidel, A., Weber, C., Ghio, M., & Bellebaum, C. (2024). Fifty Percent of the Time, Tones Come Every Time: Stronger Prediction Error Effects on Neurophysiological Sensory Attenuation for Self-generated Tones. *Journal of Cognitive Neuroscience*, *36*(10), 2067–2083. https://doi.org/10.1162/jocn_a_02226

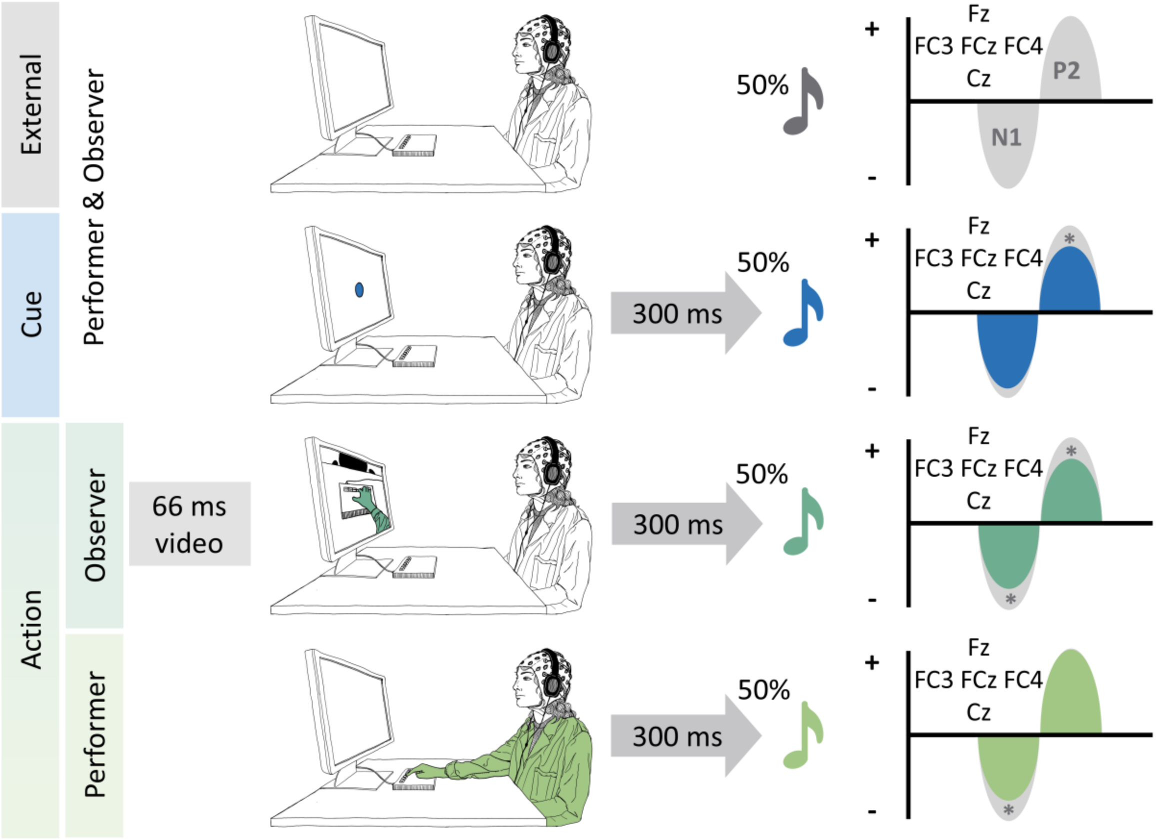
See Figure 2.3 for an overview of the action-observation setting applied in this study and the attenuation results.

2.3.1 Research Question

Sensory attenuation has been interpreted as the match of the perceived sensory input and the sensory input predicted based on the performed action (e.g., Korka et al., 2022). Previous studies have manipulated different aspects of action-outcome predictability, such as the temporal predictability and thereby when-predictions (e.g., Bäß et al., 2008), as described in Section 1.1.2, and found it to modulate attenuation relative to external tones. Few studies have examined the modulation by whether-predictions (i.e., predictions about tone occurrence), applying instead 100% action-outcome contingencies (e.g., Schafer & Marcus, 1973). Reducing the contingency to 50% through omissions, seemed not to affect the N1 (Han et al., 2022; Harrison et al., 2023), and attenuation can still be found (Seidel et al., 2021; Timm et al., 2013). However, it did affect the P2 (Han et al., 2022). As previously mentioned, the illusion of control (i.e., belief in higher [or lower] contingency) over own action outcomes also selectively affected the P2 (Seidel et al., 2021).

Interestingly, modulating contingency by altering stimulus identity (i.e., tone frequency), did show an effect on the N1: N1 attenuation was reduced for deviant tones (Knolle et al., 2013b). The P2 in that study was not analysed in this respect, because

FIGURE 2.3: ACTION-OBSERVATION SETTING AND ATTENUATION RESULTS OF STUDY 3



Note. Two groups (performers and observers) completed three conditions, in which tones were presented. External tones were temporally unpredictable in the external-tone condition (grey) and temporally predictable in the cue condition (blue). In the action conditions, the groups differed, as observers (teal) observed the button press onscreen in first-person perspective, whereas performers (green) pressed the button themselves to generate the tone. Tones followed 300 ms after button press or cue onset. Note that in action observation there was an additional interval from movement onset till button press. For precise descriptions of the conditions see Section 6.3. Importantly, tones were only presented in half of the trials. On the right, N1 and P2 amplitudes from the chosen electrode cluster are schematically depicted and significant comparisons to the external-tone condition, signifying attenuation, are marked.

deviant (i.e., mis-predicted) tones showed overlaying components that were particular to self-generation (Knolle et al., 2013b). Yet, all these studies compared high- with low-contingency conditions, not considering trial-wise variations in prediction errors.

Trial-by-trial varying prediction errors in feedback-learning studies are typically calculated based on reinforcement learning models (e.g., Sutton & Barto, 2018). The generated tone in attenuation studies does not represent a reward. Nevertheless, merely perceiving it as the consequence of one's action could be rewarding (Eitam et al., 2013) and when the contingency is reduced, participants might automatically strive to predict the uncertain action-outcomes (Clark, 2013). Predicting whether a tone follows one's action based on past tone occurrences may therefore be similar to predicting whether positive feedback follows one's action based on previous trials. In feedback-learning, action-based predictions differed from stimulus-based predictions (Kobza & Bellebaum, 2015; O'Doherty et al., 2004). Furthermore, vicarious learning, though potentially similar to active learning (Charpentier & O'Doherty, 2018), led to weaker results (Kobza & Bellebaum, 2015).

Whereas the previous two studies (Egan et al., 2023; Egan et al., 2025) and many other studies on sensory attenuation (for a review, see Horváth, 2015) implemented full (i.e., 100%) contingency between action and outcome, Study 3 explored how the prediction (errors) of tone occurrence affects N1 and P2 amplitudes in action performance and observation. For this purpose, tones were elicited by actions in only half of the trials (i.e., 50% contingency) and the prediction error was calculated for each trial based on reinforcement learning models. A group of performers and a group of observers, who saw videos of button presses likewise generating a tone in only half of the trials, were recruited. The effect of temporal predictability of the tone was also considered by including a cue condition similar to the one in Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025). While the primary research aim was to explore the modulation of N1 and P2 amplitudes by the trial-by-trial varying whether-prediction error, we first assessed whether attenuation could be found in action performance and observation, as there was no fully contingent relationship between action and outcome, which differed from the Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025). Sensory attenuation was found in previous studies even with only a 50% contingency between self-performed action and outcome (e.g., Timm et al., 2013). Thus, we expected to find a similar pattern of attenuation as in our previous studies with a 100% contingency, though P2 attenuation might be weakened (Han et al., 2022; Harrison et al., 2023). For the modulation by the whether-prediction error, we expected that smaller errors lead to smaller amplitudes (Knolle et al., 2013b), especially for action-generated tones (e.g., Kobza & Bellebaum, 2015).

2.3.2 Methods

We implemented a 2 (group) × 3 (condition) design. Action performance versus observation was a between-subjects factor in this study, as we acquired separate groups of participants who performed the button press themselves and who observed the button press via video onscreen, that is, via a picture sequence depicted in first-person perspective (see Figure 2.3). Forty participants were acquired for each of the two groups but seven did not meet the demographic inclusion criteria. The final sample comprised 36 performers and 37 observers. Next to the differentiating action condition, both groups had identical cue and un-cued external-tone conditions (see Figure 2.3). In the cue condition, external tones were preceded and made temporally predictable via visual cues. In the external-tone condition, external tones were not temporally predictable. We used a delay of 300 ms between event onset (i.e., button press or cue presentation onset) and tone throughout to ensure temporal predictability in the observer's action condition and the cue condition. The inter-trial intervals were determined either by the participant's self-timed button press (i.e., in the performer's action condition) or according to a function that generated a random interval from a normal distribution approximating the performing participant's intervals. In this study, a training block with 40 trials was included at the beginning of the experiment for action and cue conditions. The training blocks for the cue conditions and the observer's action condition were short versions of the experimental blocks and served to familiarize participants with the 50% contingency. The performer's action-condition training was also shorter than the respective experimental block. In the performer's action-condition training, button presses were visually cued so that performers learned the interval for their self-timed button presses in the experimental block. The experimental blocks each comprised 140 trials, and thus 70 tone presentations. Technically, tones were programmed into every trial but were muted in half of them. The sequence of tone and muted-tone trials was determined randomly for each block. As in the previous studies, we included control blocks to correct the ERPs of tone-unrelated activity (cf. motor correction; Horváth, 2015) for action and cue conditions, in which all tones were muted. The external-tone condition (i.e., a single experimental block) was always presented between the experimental blocks of action and cue conditions, and the respective control blocks to prevent carry-over effects on the tone expectancy. Half of the participants began the experiment with the experimental (i.e., 50% of tones muted), the other half with the control blocks (i.e., 100% of tones muted).

The pre-processing of the EEG data included a direct current detrend, filters, ocular correction independent component analysis, segmentation, baseline correction, and artifact rejection (see Section 6.3 for details). Single-trial ERPs from the experimental blocks of action and cue conditions were corrected with the averages of the respective control blocks. Based on the visual inspection of the topographies, a cluster of five frontocentral electrodes, Fz, FC3, FCz, FC4, and Cz, was chosen for analysis. For each electrode and condition, participant averages were calculated to determine the participant-specific N1 and P2 latencies. Then, N1 and P2 mean amplitudes were extracted from each exported trial in an interval of 40 ms around these latencies.

Despite the mere 50% contingency, we hypothesised, participants would attempt prediction of the occurrence of the next tone. As such, the prediction error in each trial would vary in magnitude. To calculate the prediction error, we explored three approaches. Thus, three prediction errors were calculated for each trial in action and cue conditions. The external-tone condition was not considered, as tones were not temporally predictable, therefore trials with long inter-tone intervals and trials without tone occurrence could not necessarily be distinguished, and predictions of tone occurrence could not be formed. First, the Rescorla-Wagner reinforcement learning model (Rescorla & Wagner, 1972) was applied, in which the subjective reward expectancy was the subjective expectancy of the occurrence of a tone. The prediction error was the difference between tone occurrence (tone as 1, no tone as 0) and the expectancy, which, initialised at 0.5, was defined as the addition of previous expectancy and the current prediction error weighted with the fixed learning rate of 0.2. Second, the Rescorla-Wagner model was applied with a higher learning rate of 0.7. The most recent prediction error thus had more impact on the expectancy in this calculation. Third, as comparison to the reinforcement learning models, the expectancy of the tone was determined by the frequency in the preceding 10 trials, irrespective of the prediction error.

For the attenuation analysis, the mixed-effects model had group (performer, observer) and condition (external, cue, action) as fixed effects and random effects by participant and electrode: amplitude ~ group * condition + (1 + condition | participant) + (1 | electrode). As in our previous studies, an ANOVA-like analysis was run on the models and Bonferroni-corrected pairwise comparisons were calculated where appropriate. To analyse the modulation by prediction error, fixed effects of action and cue conditions, performer and observer groups were modelled together with each of the three prediction errors: amplitude ~ group * condition * prediction error + (1 + condition

* prediction error | participant) + (1 + condition | electrode). However, the model for the P2 only included the random intercept by electrode, as the random slope for condition was removed for convergence of the model. The three prediction-error models were then compared to determine which prediction error was the best predictor for the respective ERP component.

2.3.3 Results and Discussion

A schematic summary of results can be found in Figure 2.3. Of the main and interaction effects of group and condition in the attenuation analysis, only the main effect of condition was significant for the N1, which showed attenuation in the action condition compared to the external-tone condition. Other pairwise comparisons were not significant. For the P2, both main effects and the group × condition interaction were significant. In performers, the amplitudes for the cued tones were attenuated compared to those for the un-cued external tones and to those for self-generated ones. In observers, the action and cue conditions showed attenuation compared to the external-tone condition.

The best fit for the N1 was the model with the first prediction error, namely, the one calculated with the reinforcement learning model and low learning rate. The main effects of condition and prediction error were significant as well as the three-way interaction. To resolve the interaction, separate models for performers and observers were calculated (i.e., with the same formula but excluding the group fixed effect). For the performers, both main effects and the interaction of condition and prediction error were significant. To resolve the interaction, separate models for action and cue conditions were calculated: amplitude ~ prediction error + (1 + prediction error | participant) + (1 | electrode). Prediction errors correlated negatively with the amplitude (i.e., higher error, greater N1 amplitude) in both conditions but the effect was stronger for self-generated tones. For the observers, only the main effect of prediction error was significant, providing a similar estimate as in the performer's cue condition. The best fit for the P2 was also a prediction error based on a reinforcement learning model but with the high learning rate. All main effects and interactions were significant. The resolution of the three-way interaction provided a similar pattern as for the N1. For the performers, the main effects and the interaction was significant. The prediction error correlated positively with the amplitude (i.e., higher error, greater P2 amplitude) in action and cue conditions and the effect was stronger for self-generated tones. For the observers, only the main effect of prediction error was significant, also showing the positive correlation with the P2 amplitude.

As previous studies had indicated (e.g., Timm et al., 2013), the N1 amplitudes of self-generated tones were attenuated compared to un-cued external tones, even with the 50% contingency. This was also the case for tones generated by another person, observed onscreen via picture sequence, which extends previous findings with 100% contingency (e.g., Ghio et al., 2018) to the 50% contingency setting of this study. However, the N1 in action performance and observation was only descriptively reduced compared to cued external tones, which meant that the role of action-specificity over temporal predictability in this attenuation cannot be ascertained. Previous studies showed inconsistent results (e.g., Harrison et al., 2021; Klaffehn et al., 2019). Yet, the effect of the whether-prediction error on N1 amplitudes, that smaller prediction errors led to stronger attenuation, was significantly stronger for action-generation than for cueing. This corresponds to previous findings from a study manipulating a different aspect of tone predictability, which compared tones with standard (i.e., expected) identity versus tones with deviant (i.e., unexpected) identity (Knolle et al., 2013b). Importantly, in Study 3, the stronger prediction error effect was found only in the performer group, further restricting the action-specificity to action performance. Of the prediction errors that were calculated and tested, the one that best fitted the N1 was the prediction error based on the reinforcement learning model with a low learning rate. Overall, therefore, the tone expectancy was updated slightly in each trial like the learning of response-outcome contingencies for reward maximisation. Similarly, results from reinforcement learning showed the action performance-observation distinction (e.g., Kobza & Bellebaum, 2015). While our paradigm did not include rewards as such, studies on language learning suggested that, even without rewards, brain areas of the reward system are involved in the processing of predicted input (e.g., Ripollés et al., 2014). However, in our study, the inter-tone intervals were correlated with the prediction errors and have previously been related to N1 amplitudes (SanMiguel, Todd, et al., 2013). Thus, an effect of interval length could be an alternative explanation of the overall prediction error effect (i.e., shorter intervals, lower errors, smaller N1), but not of the action-specificity.

Cued tones and tones generated by an observed button press elicited an attenuated P2 compared to un-cued external tones. Partially matching previous results in action observation (e.g., Egan et al., 2023), it thus seemed that temporal predictability would explain the P2 attenuation. Yet, the P2 attenuation of self-generated tones was not found in this study applying a 50% contingency, even though the onset of these tones was highly predictable. Considering the relation to agency judgement (Timm et al., 2016) and belief in contingency (Seidel et al., 2021), the sense of self-agency may have been reduced with the 50% contingency. The P2 in this condition in the performer group also had a higher latency, which matched the pattern of Seidel et al. (2021), who applied a 50% contingency as well, and could indicate an overlay of an attention-related component (see also Knolle et al., 2013a, 2013b). Interestingly, the effects of the prediction error were the same as for the N1: smaller prediction error, stronger attenuation; and stronger effect in action performance. The two differences to the N1 results are the pattern of attenuation and the type of prediction error. For the P2, the prediction error based on the reinforcement learning model with a high learning rate was best. The higher learning rate and thus greater impact of the most recent prediction error on the tone expectancy in the next trial could indicate a more conscious process, such as updating the believed contingency (Seidel et al., 2021) underlying the P2 attenuation.

A critical limitation of this study was the constraints of the paradigm on the estimation of prediction errors, more specifically the learning rates. In reinforcement learning, a paradigm typically includes a forced-choice task, which enables the estimations of condition- and participant-specific learning rates. A combination of this and the paradigm to investigate sensory attenuation could lead to more valid results. Furthermore, there are several options to manipulate the expectancy of tone characteristics, such as altering tone identity (e.g., Knolle et al., 2013b) instead of tone occurrence (i.e., what-predictions instead of whether-predictions), that could further be explored.

In sum, we assessed trial-by-trial varying expectancies that the participants formed regarding the occurrence of a tone as outcome of their action, when tones were only elicited in half of the trials, and the effect of the resulting prediction errors on N1 and P2 amplitudes. We tested this not just for self-generated tones (in the performer group) but also for tones generated by a human hand observed onscreen (in the observer group), in comparison to cued external tones. Regarding the research aim of assessing attenuation compared to un-cued external tones when there is no full contingency between action and outcome, N1 attenuation in action performance and observation was found but the P2 was only attenuated in action observation and for cued tones. Thus, in line with previous studies on action performance, reduced contingency seemed to affect primarily the P2 component (e.g., Han et al., 2022). There was no significant attenuation in either action performance or observation compared to cued external tones, and thus seemingly no attenuation effect beyond temporal predictability. Regarding the primary research aim, it can be said that N1 and P2 amplitudes were modulated by the trial-by-trial varying whether-prediction error. Specifically, smaller prediction errors led to smaller amplitudes, supporting the understanding of the attenuation effect as matching of perceived and predicted sensory input (e.g., Hughes et al., 2013). Interestingly, this modulation was significantly stronger in action performance for both the N1 and P2, suggesting that neurophysiological attenuation is the result of a combination of different mechanisms, including mechanisms specific to action performance. The paradigm we applied was typical for studies on sensory attenuation, which enabled comparability to previous attenuation studies but limited the precision of prediction error estimates, the improvement of which could be the aim of future research.

3 General Discussion

The focus of this dissertation was on auditory attenuation measured via N1 and P2 ERP components of the EEG in action performance and action observation. As in previous studies (for a review, see Horváth, 2015; e.g., McCarthy & Donchin, 1976), the auditory stimuli used in the three studies of this dissertation were tones, hence, environment-related action outcomes (Dogge, Custers, et al., 2019). The comparison of action performance and observation was of interest because the action observation network and its overlap to activation during action performance (e.g., Hardwick et al., 2018) marks a difference between observation and other external generation. Study 1 (Egan et al., 2023) examined the potential effect of individualism on neurophysiological attenuation in action observation, based on the findings of a perceptual study (Cao & Gross, 2015). This was done by obtaining relevant questionnaire scores and testing whether they improved the models of N1 and P2 amplitudes. Study 2 (Egan et al., 2025) investigated the interpretation of P2 attenuation as reflecting agency attribution (e.g., Timm et al., 2016), in particular as reflecting agency attribution also to another person (Ghio et al., 2021; Ghio et al., 2018). For this purpose, the action-observation conditions were extended to include a robotic agent. In Study 3 (Egan et al., 2024), a different approach was explored. Action-outcome contingency was reduced by omitting the tone in half of the trials. Prediction errors regarding whether a tone would occur were calculated (Rescorla & Wagner, 1972) and related to the N1 and P2 amplitudes to test the extent of modulation by prediction error and whether differences between action performance and observation could be found. As an overarching aim, the effect of temporal predictability on attenuation in action observation was accounted for in all three studies (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). Specifically, tones generated by an observed person were made temporally predictable by implementing a fixed action-outcome delay of at least 300 ms. Additionally, to determine the extent of attenuation that might be explained by temporal predictability, a condition was included, in which external tones were visually cued and thus temporally predictable.

The key findings can be summarised as follows: The N1 and P2 amplitudes appeared unaffected by individual differences in individualism, as well as by differences in other traits, including anomalous perception, empathy, and anthropomorphism in Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025). There was no effect of agency attribution on the P2 attenuation in Study 2 (Egan et al., 2025). In Study 3, we found attenuation without a fully contingent relationship between action and outcome (but not for the P2 of self-generated tones), and a modulation of N1 and P2 amplitudes by the trial-by-trial varying prediction error of whether-predictions (Egan et al., 2024). This prediction-error effect was stronger for self-generation than for cueing or action observation.

Irrespective of significant attenuation compared to unpredictable external tones, the N1 amplitudes of tones generated by oneself or an observed person and tones that were visually cued did not differ in Studies 1, 2, and 3 (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). For the P2, the findings diverged. For action performance and observation, there was no difference in Study 1 (Egan et al., 2023), smaller P2 amplitudes in action performance in Study 2 (Egan et al., 2025), and smaller P2 amplitudes in action observation in Study 3 (Egan et al., 2024). Furthermore, Study 3 showed differences between action performance and observation in the magnitude of the effect of whether-prediction errors on both N1 and P2 amplitudes (Egan et al., 2024). Cueing only differed from action observation in Study 1, in which the P2 in action observation was significantly reduced compared to the P2 in the cue condition (Egan et al., 2023). Taken together, these results indicate that action performance and observation do not necessarily engage the same mechanisms in processing action-generated sensory input and that attenuation in action observation does not go beyond an effect of temporal predictability with the exception of the P2 attenuation in Study 1 (Egan et al., 2023).

In the following, the results regarding the specific research questions are discussed in detail: (i) the lack of effect of individualism and other trait influences on attenuation, (ii) the potential explanation of (P2) attenuation as agency attribution, (iii) the modulation of attenuation by reduced action-outcome contingency. Furthermore, relevant aspects for all three studies, namely effects of temporal predictability and temporal proximity, are considered. Lastly, limitations and further research are addressed, and a conclusion is presented.

3.1 No Modulation by Individualism or Other Traits

In previous studies and across the three studies of this dissertation the results of perceptual and neurophysiological attenuation of tones elicited by observed actions and whether it differs from attenuation for self-generated and cued external tones was not consistent. Discrepancies in perceptual studies were related to a cultural effect (Cao & Gross, 2015). Namely, the PSE in action observation was as attenuated to temporally predictable external tones as in action performance only for a likely more collectivistic (Sato, 2008) and not for a likely more individualistic (in this case, German) sample (Weiss et al., 2011a). In the perceptual study by Cao and Gross (2015), a British sample was obtained as the individualistic group and a Chinese sample as the collectivistic group (30 participants each). Significant attenuation in action observation (i.e., when the experimenter elicited the tone) was found only for the Chinese participants. Importantly, the independent SCS scores, representing the measure of individualism, correlated with observation-attenuation values (i.e., PSE difference values) both across and within the two samples. However, Cao and Gross (2015) found no significant difference in independent SCS scores between the British and the Chinese sample. A significant difference was found only for the interdependent SCS, representing the measure for collectivism (Cao & Gross, 2015). Note that the mean scores for independent and interdependent SCS were similar in the British sample.

A modulatory effect of culture was not found on neurophysiological attenuation in Study 1 (Egan et al., 2023). Acquiring a mostly German sample, there was no significant difference in N1 or P2 amplitudes between action observation and performance, although significant attenuation relative to external tones was only found for the P2. Importantly, AICS individualism and collectivism did not improve the N1 or P2 models, indicating no significant modulation in action observation. This was despite observing larger effect sizes for neurophysiological compared to perceptual attenuation (Cao & Gross, 2015). Furthermore, Study 1 (Egan et al., 2023) aimed for a sample of 60 participants, but due to exclusions, the final sample consisted of 49 participants. Thus, although the sample was overall smaller, the sample was greater than a sub-sample of Cao and Gross (2015), for which significant correlations between SCS and attenuation values were found. In the sample of Study 1 (Egan et al., 2023), the AICS score of individualism and of collectivism, as well as the collectivism sub-scores (i.e., advice and harmony), had a mean around the medium score of the scale, similarly to the SCS scores of the British sample of Cao and Gross (2015), and participants often scored similarly high on individualism and collectivism. Regarding the individualism sub-scores, the sample's means were higher for responsibility and uniqueness, than for competitiveness. This supports the need for finer cultural differentiation as offered by the AICS sub-scores (Shulruf et al., 2007). Yet, the sub-scores did not improve the N1 and P2 models either. Overall, the variability in individualism and collectivism scores within our sample, particularly the discrepancies between the two scores, may have been insufficient to detect a relationship.

Alternatively, individualism might only affect perceptual but not neurophysiological attenuation. Few studies have measured and compared perceptual and neurophysiological attenuation, and the findings remain inconclusive (e.g., Ody et al., 2023). What Cao and Gross (2015) failed to replicate was the perceptual attenuation in action performance. This may be related to the different implementation of the visual cueing condition and the potentially increased precision of temporal predictability by the gradual transformation of the cue starting 1 s before tone onset (Cao & Gross, 2015). On the other hand, attenuation or enhancement in action performance, as described in Section 1.1.1, depended on the loudness of the stimulus relative to the participants' threshold (Reznik et al., 2015). Yet, an enhancement was also found in an online study using supra-threshold tones (Kiepe et al., 2024). They interpreted their results as attentional processes underlying perceptual attenuation (Kiepe et al., 2024). Previous findings on perceptual attenuation in the auditory modality appear to be consistent with this interpretation (Dogge, Hofman, et al., 2019; Fritz et al., 2022). As such, the cultural effect on perceptual attenuation in action observation (Cao & Gross, 2015) could also be interpreted as an inter-individual difference in the affordance of attention to another person's actions. It has also been proposed that for perceptual attenuation, heuristic judgements, which are relied upon because of the perceptual uncertainty in the paradigms, determine the effects (Reddy, 2022). The effect of prior belief about (the lack of) agency could be interpreted in this way (Desantis et al., 2012; Reddy, 2022). Accordingly, the cultural effect on loudness perception in action observation may be due to a different heuristic employed by different cultures (Cao & Gross, 2015; Sato, 2008; Weiss et al., 2011a). Further research is needed to clarify the mechanisms underlying perceptual attenuation as well as the relationship of perceptual and neurophysiological measures of attenuation.

Other individual traits that might modulate sensory attenuation in action observation are empathy and, when observing robots, anthropomorphism. As stated in Section 1.2.1, empathy has been linked to the mirror neuron system (e.g., Corradini & Antonietti, 2013; Iacoboni, 2009), but this has been criticised (Lamm & Majdandzic, 2015). Neither in the perceptual study (Cao & Gross, 2015), nor in our neurophysiological studies (Egan et al., 2023; Egan et al., 2025) was the empathy score significantly related to attenuation in action observation. Hence, empathy does not appear to be a relevant trait for altered processing of sensory input generated by another person. Study 2 (Egan et al., 2025) further included robotic agents and tested potential modulations of attenuation by anthropomorphism as well as by attitudes towards robots to account for potential bias. There was no N1 or P2 model improvement by AQ and GAToRS scores, which represent the tendency to anthropomorphise (Neave et al., 2015), and positive and negative attitudes towards robots (Koverola et al., 2022). Thus, these traits did not appear to affect the processing of tones attributed to our robotic stimuli.

Given the findings of reduced attenuation of self-generated tones in patients with schizophrenia (e.g., Ford et al., 2014; Whitford et al., 2011), previous studies have considered related traits as possible modulators of attenuation measures in action performance in healthy participants and patients with schizophrenia. For instance, Cao and Gross (2015) found a significant correlation in healthy participants between self-attenuation values (i.e., PSE difference values) and PDI scores, representing a measure of delusional ideation (Peters et al., 2004), and Heinks-Maldonado et al. (2007) found that severity scores of auditory hallucinations in patients correlated with (lacking) N1 attenuation of vocalisations relative to altered feedback. In contrast, we did not find a significant N1 or P2 model improvement by adding a measure of anomalous perception (Bell et al., 2006) in Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025). As such, anomalous perception did not appear to be a relevant trait for the degree of attenuation of sensory input generated by oneself (Egan et al., 2023; Egan et al., 2025). Although already specific to perceptual experience, perhaps a measure specific to auditory perception is needed. Supported by the findings of reduced attenuation in patients with schizophrenia and significant correlations with delusional ideation and symptom severity, symptoms, such as auditory hallucinations, have previously been related to altered forward models and resulting misattributions of agency (e.g., Feinberg, 1978; Franck et al., 2001). However, a recent meta-analysis found reduced (but not absent) N1 attenuation of vocalisation relative to a recording of it across patients and at-risk participants that was not related to symptom severity or the presence of auditory hallucinations (Mariano et al., 2024). They concluded that, rather than the cause, the reduced N1 attenuation represents a risk factor for psychosis. Studies with other patient groups have also reported diminished N1 attenuation, such as for autism spectrum disorder (auditory modality: tones; van Laarhoven et al., 2019) or obsessive-compulsive disorder (visual modality: coloured squares; Gentsch et al., 2012). Thus, the discussed findings indicate that the impairment of mechanisms yielding neurophysiological attenuation could be a risk factor for a range of disorders and potential modulators of sensory attenuation in action performance do not necessarily include traits specifically associated with schizophrenia.

3.2 Attenuation and Agency Attribution

In Study 1 (Egan et al., 2023), we found no significant differences in N1 or P2 amplitudes between action observation and performance. This stood in contrast to previous perceptual findings (Weiss et al., 2011a, 2011b; Weiss & Schutz-Bosbach, 2012) and to previous findings for the N1 (Ghio et al., 2021) and the P2 (Ghio et al., 2018) component. It was also in contrast to the P2 results of Study 2 (Egan et al., 2025), which showed attenuation in action observation relative to external tones, but stronger attenuation for self-generated tones. These inconsistencies in findings might be due to different methodological implementations of the action-observation condition, such as perspective, experimental design (i.e., within-subject or between-subjects), and setting.

Whereas action observation is typically experienced in a third-person perspective, as it was in the perceptual studies (e.g., Weiss et al., 2011a) and in Ghio et al. (2021), action observation was implemented in a first-person perspective in Ghio et al. (2018) and Studies 1, 2, and 3 (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). Interestingly, view-dependent mirror neurons were found (with a slight preference for first-person perspective) in monkeys, which were proposed to facilitate action perception (Rizzolatti & Fogassi, 2014). In humans, studies found processing advantages for the first-person perspective as opposed to the third-person perspective in EEG (e.g., earlier activation of motor-premotor regions; Angelini et al., 2021; Angelini et al., 2018) and also stronger activation of frontoparietal (mirror-neuron) and visual regions in fMRI (Ge et al., 2018). Seidel et al. (2023) examined attenuation relative to external tones during action-observation in first- and third-person perspective, also considering the temporal development by accounting for effects of first and second block, and of trials within a block. The N1 was attenuated in action observation, but for the first-person perspective, the attenuation developed only in later trials (for both blocks; Seidel et al., 2023). This was related to the perspective being unfamiliar in action observation, which could have hindered prediction. This may be a critical point for the action-observation conditions in Studies 1, 2, and 3 (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). However, in these studies, no significant differences to the N1 amplitudes of self-generated tones were observed. The familiarisation with the first-person perspective may have been facilitated, as trial numbers in an experimental block were larger (i.e., > 50) and, in Studies 2 and 3 (Egan et al., 2024; Egan et al., 2025), a single block for each condition was applied. In Seidel et al. (2023), the P2, on the other hand, was attenuated irrespective of perspective. Therefore, the perspective of the action observation does not appear to explain the discrepancies across findings of these and previous studies.

Some previous studies compared action performance and observation as either a within-subject or a between-subjects factor. In other words, participants faced the action-outcome association by performing the actions themselves and by observing another person (e.g., Weiss et al., 2011a), or participants were assigned to be either the performer or the observer, as was done in both previous neurophysiological studies of interest (Ghio et al., 2021; Ghio et al., 2018). The latter could be affected by inter-individual differences, while lacking the direct learning experiences and potential carry-over or order effects of the former. In an fMRI study with professional ballet dancers, observation of moves that the observers had motor experience with showed increased activation of premotor, parietal, and cerebellar regions (Calvo-Merino et al., 2006). However, the experience of pressing a specific button might be less relevant than that of performing more complex ballet moves. Indeed, comparing studies with similar action-observation setting but different experimental designs, Ghio et al. (2018) and Study 2 (Egan et al., 2025) showed no differences in N1 or P2 attenuation, whereas Ghio et al. (2021) and Study 1 (Egan et al., 2023) differed in the N1 results. Yet, rather than due to motor experience, this seemed to be due to the introduction of an action-outcome delay in Study 1 (Egan et al., 2023), which made the tones generated by the other participant more predictable in time. Therefore, the differences in design do not appear to explain the discrepancies across findings either.

Focusing on the discrepancies in P2 attenuation in action observation, what differed between the studies was the action-observation setting, which was either live or via video. Interestingly, the P2 in action observation in video settings was less attenuated than for self-generation (Egan et al., 2025; Ghio et al., 2018), and not different to cued external tones (Egan et al., 2024; Egan et al., 2025). In contrast, Ghio et al. (2021) found, similarly to Study 1 (Egan et al., 2023), that the P2 in live action observation was as attenuated as in action performance. A difference between these settings, which might affect the activation in the action observation network, was that in the videos only the hand (and forearm) was visible, while during live observation the participant viewed or could view more of the person they observed. In an fMRI study testing two types of view during action observation, Turella et al. (2009) found no differences in activation of the action observation network between observing videos of grasping motions performed by a person or by an isolated hand. Yet, findings from motor imagery tasks with EEG indicate differences between observing hand-clenches in a video and observing hand-clenches simulated by a robotic glove in front of the participants (Tang et al., 2025). Potential explanations for the stronger P2 attenuation when observing a person live could relate to the sense of agency. This agency-based interpretation of the P2 attenuation was suggested by previous studies (e.g., Knolle et al., 2013a; Seidel et al., 2021; Timm et al., 2016) and was also discussed for action observation (Ghio et al., 2021; Ghio et al., 2018). As such, the vicarious sense of agency might be increased in live observation because of the possibility of interaction, even though it was not included in the task, and social interaction during EEG preparation (taking at least 40 min) or even as previous acquaintances in some cases.

However, we attempted to test the agency-based interpretation of the P2 attenuation in Study 2 and found that the P2 was not affected by different agent scenarios in video observation (Egan et al., 2025). We noted increased agency judgement inaccuracies and decreased agency-confidence ratings for the action-observation conditions. It could be that this decrease of agency led to lesser P2 attenuation. Yet, agency-confidence ratings and agency judgement accuracy did not significantly improve the (N1 or) P2 model. On the one hand, Weller et al. (2017) also found diverging effects of filling the action-outcome delay on the rated agency and the P2. On the other hand, as observed for delays, effects on explicit agency may depend on the alternatives provided (e.g., self-generated, self-generated but altered, and externally generated) and the wording of agency questions (Wen, 2019).

Importantly, (video) action-observation conditions with both human and robotic agents as well as the cue condition did not differ in P2 amplitudes. A possible explanation comes from studies on intentional binding, a phenomenon that has also been interpreted as an implicit measure of the sense of agency (Hughes et al., 2013). Binding describes a perceived timing shift: The action is perceived to occur later in time and the outcome (e.g., tone after 250 ms) is perceived to occur earlier in time (Haggard et al., 2002). Originally binding was measured by presenting a clock face with a single clock hand (i.e., Libet clock) to enable the tracking of time and by giving instructions to judge when either the action or the outcome had occurred (Haggard et al., 2002). Binding can also be found with an interval-estimation task (i.e., participants underestimate the interval that passed between action and outcome) without the distinction of action and outcome shift (for a review, see Moore & Obhi, 2012). Studies applying this paradigm have shown vicarious sense of agency, that is, the binding effect in action observation beyond mere visual cueing similar as for action performance (Poonian & Cunnington, 2013; Poonian et al., 2015). With the Libet clock task, studies further found that the action shift for live action observation of robotic agents depended on human-likeness and perceived intentionality (Roselli et al., 2022; Roselli et al., 2021). Therefore, binding also depended on whether the observers adopted the intentional stance (i.e., the observed behaviour follows intentions) or the design stance (i.e., the observed behaviour is designed) towards the robotic agent (Dennett, 1971). Similarly to binding, intentionality might be relevant for P2 attenuation and the videos used in Study 2 (Egan et al., 2025) might have rather led to participants taking the design stance towards the depicted human and robotic agents to predict button presses, instead of ascribing intentions.

Given the speculated differences between live and video observation, these agent scenarios could be implemented in a live action-observation setting to investigate whether the P2 attenuation is affected by agency attribution depending on the setting. Alternatively, they might be implemented in virtual reality (VR). In a modified attenuation paradigm using VR, the external-tone condition was implemented with a robotic agent (Feder et al., 2023). Tones were produced during the use of a keypad to open doors in VR and were delayed by various fixed delays. The robotic agent was a sphere, thus lacking human likeness. Furthermore, tones generated by the robotic sphere were not temporally predictable. The N1 amplitudes were attenuated across delay conditions relative to the robot generation, whereas the P2 amplitudes were not attenuated. While it is difficult to compare the results of this study to those of studies applying the contingent paradigm, it does provide a future direction to investigate various agent scenarios and their effect on the P2 in a way that is more controlled than live observation but more immersive than video observation. Indeed, it appears that the intentional stance is adopted for onscreen VR characters believed to be human-controlled and affects ERP components later than 170 ms (Caruana & McArthur, 2019).

As an alternative interpretation of the P2 results, the social context of the acquisition in live action observation settings may have differential effects on attention and arousal, which have been shown to increase N1 and decrease P2 amplitudes (Crowley & Colrain, 2004). Therefore, in comparing the live action observation in Study 1 (Egan et al., 2023) and the video action observation in Study 2 (Egan et al., 2025), the N1 attenuation compared to external tones that only reached significance in Study 2 and the P2 attenuation compared to cued external tones that was evident only in Study 1 might alternatively be explained by an overlaying effect of increased attentiveness and arousal during live observation. Then, in a live setting or in immersive VR, the P2 in action observation might be as attenuated as for action performance but would not depend on the agent who is observed.

In sum, the inconsistent results of previous studies on neurophysiological attenuation in action observation do not appear to be due to different implementations of perspective or design. For the P2 component, however, the action-observation setting seems to be a relevant factor for the degree of attenuation (i.e., stronger for live observation). While this might relate to the perceived sense of agency, we did not find an effect of agency attribution on the P2 in video observation (Egan et al., 2025). However, participants might have adopted the design stance instead of the intentional stance (Dennett, 1971) in this setting. Future research could investigate whether agency attribution effects on the P2 are evident when participants take the intentional stance towards the observed agents by applying live observation or observation in VR.

3.3 Prediction Under Action-Outcome Uncertainty

Previous interpretations of sensory attenuation build upon the contingent relationship between action and tone that enables accurate predictions and might also in this way enable the sense of agency (Horváth, 2015; Kiepe et al., 2021). However, even without a contingent relationship between button press and tone, the N1 amplitudes of tones decreased linearly with increasing temporal proximity to the button press (e.g., Horváth et al., 2012). Furthermore, for self-generated tones with uncertain occurrence (i.e., 50% contingency), N1 attenuation could be observed in previous studies (e.g., Seidel et al., 2021; Timm et al., 2013), as we did in Study 3 (Egan et al., 2024). This was not the case for the P2 in Study 3 (Egan et al., 2024). Interestingly, Seidel et al. (2021) did not find a P2 attenuation either, but an illusion-of-control effect on the P2 that might be interpreted as an effect of the belief in contingency and that was specific to self-generation (i.e., not found for external tones). More precisely, P2 amplitudes were reduced when a belief in high contingency was induced compared to when a belief in low contingency was induced. This was further shown to be the case particularly for participants with higher differences in control ratings (i.e., more effective illusion of control). Applying a different paradigm, in which a time frame was given for the (in)action, Han et al. (2022) investigated the effect of contingency by comparing tones produced with 100% and 50% action-outcome or inaction-outcome association. Tones were temporally predictable throughout. They found reduced amplitudes in the 100% condition for the P2 but not for the N1 and concluded that the former reflects contingency, which seems to be in line with the results of Seidel et al. (2021). Similarly, Harrison et al. (2023) compared 100% and 50% action-effect contingency with further investigating differences between stimulus-driven and volitional actions, and they found no significant effect of contingency on the N1. The P2 was not analysed but the graphical depiction of the ERPs suggested no difference between 100% and 50% contingent tones as elicited by volitional action. Taken together, unlike the N1, the P2 appeared more sensitive to contingency (i.e., whether-prediction) in some studies.

Considering the broader ecological contingency of action-outcome associations, studies have investigated the effect of the type of action eliciting the auditory stimulus. van Elk et al. (2014) compared hand- and foot-generated tones (i.e., pressing the button with the big toe), finding no difference in N1 attenuation and stronger P2 attenuation for hand-generated tones (for foot-generation no attenuation was found). Mifsud, Beesley, et al. (2016), on the other hand, first compared button presses and saccades as eliciting actions and observed P2 attenuation only for button presses as well as stronger N1 attenuation (see Mifsud et al., 2018, for stronger N1 attenuation of visual outcomes for saccades rather than button presses). By extending the types of action with soundless blowing, Mifsud and Whitford (2017) aimed to include a more common association, namely, mouth-actions and auditory outcomes without the confounds of vocalisation. They showed that N1 and P2 attenuation was strongest for the mouth and hand-related actions, though amplitudes were descriptively smallest for mouth-related ones, especially for the P2 (Mifsud & Whitford, 2017). They interpreted this as the relevance of habitual associations (i.e., higher prevalence in daily life) for attenuation, particularly for attenuation of the P2. Ross et al. (2017) further implemented an ecologically valid action-outcome association by using a singing bowl. They found stronger N1 attenuation for bowl-actions than producing the same sound by button press and descriptive P2 attenuation only for bowl-actions. An external-tone condition was included before and after the self-generation condition. While the N1 amplitudes decreased within each external condition, the P2 was larger in the external condition after self-generation only when bowl-actions were performed. Therefore, the findings suggested that prior experience with ecologically valid action-outcome associations can modulate (i.e., increase) the P2 amplitude of sounds that are identical to the outcome sounds. Taken together, the P2 seems more sensitive to contingency also in terms of habitual associations, which can be interpreted to be in line with the proposed relevance of agency (e.g., Timm et al., 2016) and intentionality (Egan et al., 2025).

Study 3 (Egan et al., 2024) investigated, beyond the effect of the overall 50% contingency on attenuation, also the potentially modulating effect of trail-by-trial varying prediction errors on the N1 and P2 amplitudes. Indeed, for both the N1 and the P2 there was a significant effect of whether-prediction errors: The smaller the error, the smaller the amplitude (i.e., in case of the N1, the stronger the attenuation). Interestingly, for both, the best fit was a prediction error calculated based on reinforcement learning, that is, the expectancy of tone occurrence in each trial resulted from the integration of the previous expectancy and the most recent (weighted) prediction error. This supports the assumption that participants attempted to predict the occurrence of tones based on their previous experience, even though it was not part of the task (Clark, 2013), and that prediction errors across different contexts of associative learning are conceptually similar (Sutton & Barto, 2018). For instance, intentional binding (specifically, the action shift) also appeared affected by prediction-error learning (Moore & Haggard, 2008; Moore et al., 2011). Moreover, correct predictions were processed similarly as rewards in studies on language learning (Orpella et al., 2021; Ripollés et al., 2014). However, whether the reward system is also involved in predicting tone occurrence, as in Study 3 (Egan et al., 2024) should be addressed in future research.

That ERP components reflect the prediction error has been suggested in particular for the N1 (Schröger et al., 2015). Previous studies have analysed the ERP components in omission trials (for a review, see Korka et al., 2022). Omission N1 and N2 were found only for rare omissions (i.e., 12% not 50%) and interpreted as sensory template of prediction turned prediction error (SanMiguel, Widmann, et al., 2013). In a study analysing both omission and tone trials, Tast et al. (2024) showed the differential effects of rare (i.e., 20%) and frequent (i.e., 50%) omissions on the omission N1 and the N1 attenuation, which were consistent with the contingency sensitivity of the omission N1 (SanMiguel, Widmann, et al., 2013) and the contingency insensitivity of the N1 attenuation in Study 3 (Egan et al., 2024). They noted this as support for alternative interpretations for N1 attenuation, such as sensorimotor gating (i.e., attenuated processing during movement; e.g., Schröger et al., 2015), which is quasi-predictive, in that it reflects environmental probabilities without tracking them (Press et al., 2023). Yet, the prediction-error effect on the N1 that we found in Study 3, especially for self-generation (Egan et al., 2024) stands in contrast to this, as well as several studies showing reduced N1 amplitudes of tones with predictable identity (i.e., frequency) relative to tones with mis-predicted (or unpredictable) identity (e.g., Bäß et al., 2008; Knolle et al., 2013b; Kühn et al., 2011; but see McCarthy & Donchin, 1976, for attenuation only of predicted versus mis-predicted external tones). An overall identity-prediction effect was found also when including self-timed and cued self-generation, temporally predictable and unpredictable external tones (Harrison et al., 2025). This, in turn, is in line with specific prediction yielding N1 attenuation. However, this was not always replicated (e.g., Dogge, Hofman, et al., 2019). In particular, Tast et al. (2025) found an effect of identity prediction only on the N1 omission but not on the N1 attenuation. Taken together, specific and unspecific mechanisms seem to modulate the N1 time range of self-generated tones (cf. Tast et al., 2025).

Similarly, the P2 results in Study 3 (Egan et al., 2024) could suggest multiple underlying processes: the prediction-error effect attenuating the P2 and a general enhancement (i.e., lack of attenuation) of the P2 in self-generation with the 50% contingency. On the one hand, the prediction-error effect may reflect trial-by-trial updating of the believed contingency (see also Seidel et al., 2021). That the prediction error based on the high learning rate was the better fit for the P2 might be interpreted as a more conscious underlying process. On the other hand, the P2 of self-generated tones for zero or small prediction errors was still visibly enhanced compared to cued tones, which could suggest a separate underlying mechanism. In a recent meta-analysis of neuroimaging studies on sensory attenuation across various modalities, a cerebellar co-activation network was identified as one of two that drive attenuation, supporting the role of the cerebellum for the processing of self-generated sensory input (Gu et al., 2025). The second network, centred in a temporal region, showed stronger activation for external than self-generated input and was interpreted in terms of sensory conflict detection and reduced agency. Because the low contingency between action and tone in Study 3 (Egan et al., 2024) likely reduced the sense of agency (see also Seidel et al., 2021), this second network might relate to the enhancement (i.e., the lack of P2 attenuation of self-generated tones). However, this enhancement was specific to self-generation (i.e., P2 attenuation was found in action-observation and cue conditions) as was the stronger prediction-error effect. Specifically, for both N1 and P2 amplitudes, the prediction-error effect was stronger for self-generation compared to the cue condition, while for observers there was no difference between the prediction-error effect in action observation and cueing (Egan et al., 2024). For the self-specific P2 enhancement, an overlay of a later positivity due to the uncertainty of tone occurrence and greater saliency of own action-outcomes (Knolle et al., 2013a, 2013b; Seidel et al., 2021) might be considered as an explanation. The stronger prediction-error effect on N1 and P2 amplitudes for self-generation was further in line with results from feedback learning (Bellebaum et al., 2012; Bellebaum et al., 2010; Kobza & Bellebaum, 2015) and might similarly be related to greater saliency of own action outcomes.

An aspect that was listed as one of the confounds by Hughes et al. (2013) and was exclusive to the self-generation condition compared to all others in Studies 1, 2, and 3 (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025) was the control over the temporal onset of the tone. If the effect of whether-prediction errors on N1 and P2 is partly linked to temporal control, then it might be interpreted as an increased feeling of control (and more attenuation) when the tone occurrence was more expected. However, findings on the effect of temporal control on N1 and P2 amplitudes are inconsistent (Harrison et al., 2021; Kaiser & Schütz-Bosbach, 2018; Sowman et al., 2012; Timm et al., 2016), as are such findings on PSE values (Kiepe et al., 2024; Weiss et al., 2011b). Overall, a feeling of lacking control might yet relate to the contingency effect on the P2 and thus the lack of attenuation that we found for self-generation (Egan et al., 2024). Another line of studies has focused on the motivational aspect of control, and thus contingency, highlighting the connection to reinforcement learning, in that enacting control is rewarding in itself (Eitam et al., 2013). For instance, when participants randomly selected buttons associated with different outcome probabilities (e.g., low: 0%, and high: 90%), Penton et al. (2018) showed that response times were affected by the probabilities only when contingency was high (i.e., low probabilities of the stimulus occurring independently of the button press). Thus, the effects of action-outcome uncertainty that are specific to self-generation might relate to the temporal control over one's own action outcomes and their higher saliency or motivational incentive.

3.4 Temporal Predictability and Temporal Proximity

As previously mentioned, the contingent paradigm has been criticised for containing confounds in the comparison of self-generated and external tones (Hughes et al., 2013). One of them was temporal predictability, in that the temporal onset of the tone can be predicted only for self-generation but not for external tones. Consequently, the attenuation of N1 and P2 amplitudes might reflect general predictive mechanisms. This alternative explanation would also apply for the attenuation in action observation. In action observation, the temporal proximity of action and tone becomes relevant, as it can determine the temporal predictability of the tones generated by observed actions (e.g., Ghio et al., 2021). Moreover, the effect of temporal proximity was raised as a separate issue for N1 and P2 attenuation of self-generated tones (e.g., Horváth et al., 2012).

3.4.1 Action Observation Provides Temporal Predictability

Temporal predictability was not accounted for in the action-observation studies by Ghio et al. (2018) and Ghio et al. (2021), but in Studies 1, 2, and 3 (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025) via the increased action-outcome delay and the cue condition. We found no significant difference between action observation and cueing in N1 and P2 amplitudes, and the whether-prediction error effect in Studies 2 and 3 (Egan et al., 2024; Egan et al., 2025). Regarding the N1 amplitudes, this was in line with Study 1 (Egan et al., 2023) and with the interpretation of the N1 attenuation as reflecting temporal predictability, which was evident in the comparison of N1 results for action observation with and without an action-outcome delay (Egan et al., 2023; Ghio et al., 2021; Ghio et al., 2018). A recent fMRI study investigated differences between conditions with and without actions, with and without tones, for action performance and observation (Ghio et al., 2025). Action observation was implemented as a video, thus matching the action-observation setting of Studies 2 and 3 (Egan et al., 2024; Egan et al., 2025), for which we found no differences in N1 and P2 amplitudes to cued tones. Ghio et al. (2025) found that tones generated by self-performed and observed actions as well as external tones yielded differential activation in the right auditory cortex. In an exploratory analysis, the activation was only descriptively attenuated in action performance and observation. Regarding their analysis of other regions, including cerebellar regions, their results were in line with the postulated cerebellar forward model and its efference-copy input for self-generation. Yet, predictive mechanisms in action observation could not be related to the cerebellum or motor regions. This contrasts with findings of cerebellar activation as part of the extended action observation network (for a review, see Antonioni et al., 2024). Still, the results of Ghio et al. (2025) as well as findings of Studies 2 and 3 support the notion that video action observation involves more general predictive mechanisms as would be the case for other temporally predictable external tones (e.g., Egan et al., 2024; Egan et al., 2025) and the notion that the measured attenuation is the result of a combination of different predictive mechanisms (e.g., Egan et al., 2024).

3.4.2 The N1 Component

Importantly, in Studies 1 and 2 (and 3), there were also no significant differences in N1 amplitudes of self-generated tones compared to cued tones or tones generated by observed button presses (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025), while the N1 of cued tones (Egan et al., 2023) or of tones generated by an observed human agent (Egan et al., 2025) were attenuated relative to unpredictable external tones. This could be interpreted as tentative support for previous interpretations of N1 attenuation as reflecting temporal predictability also for self-generation (e.g., Harrison et al., 2021). However, three aspects warrant further discussion: (i) the lack of attenuation for self-generation relative to unpredictable external tones in Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025), (ii) the inconsistent findings of previous studies regarding the comparison of self-generated and cued external tones (e.g., Harrison et al., 2021; Klaffehn et al., 2019), and (iii) the findings of previous studies regarding the comparison of predictable and unpredictable temporal onsets that do not support the relevance of temporal predictability for N1 amplitudes of self-generated tones (Bäß et al., 2008; Lange, 2011).

Notably, in Studies 1 and 2 the N1 of self-generated tones was not significantly (i.e., only descriptively) attenuated relative to unpredictable external tones (Egan et al., 2023; Egan et al., 2025). This was found despite the use of relatively long inter-tone intervals (i.e., 2-3 s), for which N1 attenuation was previously found to be larger (SanMiguel, Todd, et al., 2013). Yet, studies applying action-outcome delays for self-generated tones have observed learning effects with N1 attenuation developing over the first 100 trials (e.g., Aliu et al., 2009; Cao et al., 2017; Elijah et al., 2018). Thus, one potential reason for the lack of N1 attenuation could be the implemented action-outcome delay (i.e., ≥ 300 ms), although it was consistent throughout the experiment and, considering both training and experimental blocks, the number of trials exceeded 100. Still, this might have been insufficient (i.e., requiring at least 100 training trials), perhaps particularly for the length of the delay, which was larger than of the delay primarily tested in the studies on learning effects (i.e., 100 ms; e.g., Aliu et al., 2009). In studies applying fixed delays over 100 ms, the N1 was attenuated more without the delay (Aliu et al., 2009) or as much (Klaffehn et al., 2019; Timm et al., 2016), but attenuated even for delays of 750 ms with trial numbers comparable to those in our studies (Klaffehn et al., 2019; van Elk et al., 2014). However, in the study by van Elk et al. (2014) the comparisons of the conditions with no or shorter delays (i.e., 0, 250, and 500 ms) to external tones were not significant either (see also Whitford et al., 2011, for N1 attenuation of button-press-elicited vocalisation only without delay). Taken together, effects of temporal proximity on the N1 have so far been inconsistent but some findings indicated that action outcomes with diminished temporal proximity require more trials to show N1 attenuation. Furthermore, the effect of inter-tone intervals was previously mentioned as a limitation for Study 2 (Egan et al., 2025). Because increased intervals increase the N1 (and P2) amplitudes (Crowley & Colrain, 2004; Pereira et al., 2014; SanMiguel, Todd, et al., 2013), the interval differences between conditions could be an alternative explanation for the lack of N1 attenuation for self-generation in Study 2 (Egan et al., 2025). However, in this study the intervals were largest for action-observation and cue conditions, for two of which significant N1 attenuation was yet found. As possible additional analyses when using the single-trial approach and if inter-tone intervals can be extracted for the trials included in the main analysis, inter-tone intervals could be added as fixed effect in linear mixed-effects models and thus, the effect of intervals could be accounted for (for an example, see Bolt & Loehr, 2021). Moreover, trial number could serve as another fixed effect in the model to examine the temporal progression of attenuation effects (e.g., Seidel et al., 2023), such as differential learning effects for delayed action-outcomes.

As described in the introduction, studies examining the effect of temporal predictability on the N1 by comparing self-generated and cued tones, have reported mixed results (e.g., Harrison et al., 2021; Klaffehn et al., 2019). Temporally predictable external tones, despite various implementations (e.g., highly regular intervals, countdowns, and hand-clap videos), showed descriptive (Sowman et al., 2012) and significant N1 attenuation relative to unpredictable external tones (Harrison et al., 2025; Harrison et al., 2021; Kaiser & Schütz-Bosbach, 2018; Schafer & Marcus, 1973; van Laarhoven et al., 2021; Vroomen & Stekelenburg, 2010), as we found in Study 1 (Egan et al., 2023). However, only few studies found no significant difference in N1 amplitudes between temporally predictable external tones and self-generated ones (Harrison et al., 2025; Harrison et al., 2021; Sowman et al., 2012), as we found in all three studies (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). Considering cued external and cued self-generated tones (i.e., not self-timed), Kaiser and Schütz-Bosbach (2018) found an enhancement for the latter, as opposed to Sowman et al. (2012) who found no significant difference, and to Harrison et al. (2021), who found stronger attenuation for these (perhaps related to attentional suppression, see Harrison et al., 2023). Several studies, on the other hand, found an attenuation of (self-timed) self-generated tones compared to temporally predictable external ones (Klaffehn et al., 2019; Lange, 2011; Mifsud, Oestreich, et al., 2016; Schafer & Marcus, 1973; Timm et al., 2016). In those studies, intervals of 1000 ms (Mifsud, Oestreich, et al., 2016) and 2000 ms (Schafer & Marcus, 1973) to predict the tone onset were perhaps too long to form highly precise temporal predictions compared to intervals or movement-videos of 500-ms duration (e.g., Kaiser & Schütz-Bosbach, 2018; van Laarhoven et al., 2021). Yet, the 750-ms delay in the study by Klaffehn et al. (2019) offered precise prediction of temporal onset due to the implementation of a loading bar. For comparability, a visual stimulus moved back and forth between trials. Participants might have timed their button presses accordingly, motivated by their displayed mean interval after 25 trials, and thus cueing their actions, which can lead to further attenuation (Harrison et al., 2021). Overall, the presence or absence of N1 attenuation of self-generated compared to cued tones appears to depend on the implementation of the cueing, particularly in terms of the prediction precision relative to the highly precise temporal prediction based on one’s own action.

Previous studies have also investigated the effect of temporal predictability on N1 attenuation by comparing predictable and unpredictable temporal onsets. Studies on self-generated tones showed no effect of temporal predictability (Bäß et al., 2008; Lange, 2011), whereas a study on pitch-shifts during vocalisation (Behroozmand et al., 2016), and studies investigating predictable external tones and other sounds did (Lange, 2009; van Laarhoven et al., 2021; Vroomen & Stekelenburg, 2010). Regarding temporal predictability, the applied temporal range (i.e., difference between minimal and maximal delay) and the probability of a given delay might be considered as separate aspects (cf. Tanaka et al., 2019). While the temporal ranges in the above-mentioned studies were mostly comparable (480-570 ms, but 800 ms in Lange, 2011), some ranges were located at different points in time: In particular the studies including self-generation placed the temporal range after the event (e.g., with a maximal delay of 1000 ms; Bäß et al., 2008; Behroozmand et al., 2016); the studies focused on visual cues placed the temporal range around the event (i.e., including trials with auditory stimuli presented before the event; van Laarhoven et al., 2021; Vroomen & Stekelenburg, 2010). For the latter, the average delay in the experiment was, therefore, around 0 ms. Moreover, in a study with a probable immediate tone (probability = 0.7), the N1 of self-generated tones was also attenuated for small delays of 50 and 100 ms (probability = 0.1, each) but not for a delay of 250 ms (Pinheiro et al., 2019). This tolerance for small temporal deviations resembles the tolerance for various deviations found particularly when action outcomes has been learned with a delay (Aliu et al., 2009; see also Elijah et al., 2018). Taken together, a greater average delay might lead to higher tolerance for temporal onset variability and abolish the effect of temporal predictability on N1 amplitudes that is found for average delays of 0 ms (particularly with large temporal ranges and low probabilities in the unpredictable-onset condition). Alternatively, N1 attenuation of self-generated tones might rely on other mechanisms than temporal predictability.

It should be noted that the N1 attenuation patterns in all three studies did not neatly align with the temporal-predictability interpretation. Indeed, the N1 amplitudes of cued external tones in Studies 2 and 3 were descriptively closer to the external tones than self-generated ones (Egan et al., 2024; Egan et al., 2025). All in all, the effect of temporal predictability on N1 amplitudes appears to not be robust and to depend on the temporal proximity of the outcome to the event and resulting deviation tolerance, as well as on specific implementations of cue conditions (e.g., regarding the precision of temporal prediction). For delayed event-outcomes, the degree of N1 attenuation may further depend on the number of learning trails. Future studies could clarify whether N1 attenuation for self-generation (relative to unpredictable external tones) relies on specific mechanisms, as suggested by findings regarding the cerebellum (e.g., Ghio et al., 2025) and the prediction-error effect in Study 3 (Egan et al., 2024), that provide higher precision of temporal predictions and weigh the temporal predictability of the outcome differently than predictive mechanisms for temporally predictable external tones.

3.4.3 The P2 Component

The P2 seemed to in part reflect temporal predictability as well, because the P2 of cued tones was attenuated relative to unpredictable external tones in all three studies (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). Significant P2 attenuation for temporally predictable external tones (compared to unpredictable ones) was also found in previous studies with similar (500 to approx. 692 ms; Sowman et al., 2012; van Laarhoven et al., 2021) and much longer (1000-2000 ms; Mifsud, Oestreich, et al., 2016; Schafer & Marcus, 1973) intervals between (last) cue, previous tone, or movement onset and the tone. An effect of temporal predictability was also evident in the comparison of predictable and unpredictable temporal onset (van Laarhoven et al., 2021; Vroomen & Stekelenburg, 2010). However, cued external tones did not show P2 attenuation relative to unpredictable external tones in the study of Harrison et al. (2021). They further did not find P2 attenuation in a self-generation condition, namely the cued (opposed to self-timed) self-generation (Harrison et al., 2021). Thus, in contrast to the above-mentioned studies, temporal predictability was not sufficient to yield P2 attenuation. They and other studies also found significant P2 attenuation for self-timed self-generated tones compared to temporally predictable external tones (Harrison et al., 2021; Klaffehn et al., 2019; Mifsud, Oestreich, et al., 2016; see also Timm et al., 2016, though not self-timed), as we did with an action-outcome delay in Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025). Hence, P2 attenuation has been related to temporal control and sense of agency, and we explored the effect of agency attribution in Study 2 (Egan et al., 2025).

Another study did not observe P2 attenuation relative to external tones for cued tones as well as for a self-generation condition, namely the delayed (opposed to immediate) self-generation (Klaffehn et al., 2019). P2 attention for (immediate) self-generated tones, though not always analysed (Horváth, 2015), has been found consistently (e.g., Klaffehn et al., 2019; Schafer & Marcus, 1973) across various inter-tone intervals (SanMiguel, Todd, et al., 2013). It was also found with a delay of 100 ms but, unlike for the N1, without an increase in attenuation over time (Cao et al., 2017). Yet, findings of P2 attenuation with delays and thus decreased temporal proximity between action and tone have been mixed (e.g., Klaffehn et al., 2019, with a 750-ms delay). For instance, van Elk et al. (2014) found P2 attenuation for 0 and 750 ms delays (and increase in amplitude for 250 and 500 ms), and Pinheiro et al. (2019) found significant attenuation only for the probable immediate tones, while reporting an increase in amplitude with increase of delay (50, 100, and 250 ms). Implementing a cued self-generation condition, Timm et al. (2016) presented a test tone immediately or after 200 ms, which followed an adaptation phase with 0, 200, or 400 ms (i.e., the delay was as adapted to or shorter). While a main effect of delay on the P2 (i.e., greater amplitudes with delay) was observed, the effect was stronger for self-generation. Specifically, for delayed self-generated tones, they found an enhancement compared to their comparison condition, irrespective of delay continuity or deviation (Timm et al., 2016). Taken together, action-outcome delays can increase P2 amplitudes and the temporal range applied in the experiment, which was larger in the above-mentioned studies (i.e., > 0) than in Cao et al. (2017) and Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025), might modulate the P2 attenuation for delayed self-generated tones. Namely, only when participants face various action-outcome delays, instead of a consistent temporal relation between action and outcome, does temporal proximity seem relevant for P2 attenuation relative to external tones. This was also observed in the VR study, in which participants operated virtual keypads (Feder et al., 2023). They applied a variety of fixed delays in separate blocks and P2 amplitudes increased with increasing delay, showing enhancement relative to external tones. The P2 enhancement for delayed tones, when participants have experienced higher temporal proximity between action and outcome, could be interpreted in terms of decreased perception of the tones as action outcomes. However, in the study by Feder et al. (2023) participants could not detect larger delays (e.g., 300 ms) more often. As such, they likely retained a sense of agency for the delayed tones, despite the P2 enhancement.

3.4.4 Attention To Cues and Button Presses

Previous literature has pointed out that attentional effects may be alternative explanations for attenuation or have additional effects on N1 (and P2) amplitudes (Horváth, 2015; Schröger et al., 2015): The N1 is decreased with decreased attention to the tone (Näätänen & Picton, 1987), and the action to elicit the tone is assumed to draw attention away from it. This could also be an alternative explanation for attenuation for cued external tones, as the cue could also draw attention away from the following tone. In this case, the N1 was attenuated relative to external tones because the cue drew more attenuation in Study 1 (Egan et al., 2023), possibly due to the lesser familiarity, the button press by human agents drew more attention in Study 2 (Egan et al., 2025), possibly due to social salience, and the button press of oneself or another drew more attention in Study 3 (Egan et al., 2024), possibly due to the action-outcome uncertainty. One study investigated the effect of attention by implementing an external-tone condition with an additional task that made the tones task-relevant (i.e., detection of deviant intervals) and found overlapping but distinct attenuation and attention effects on the N1 (Saupe et al., 2013). Alternatively, participants had to count the occurrence of a particular event (e.g., tones as action-outcomes), thereby directing their attention to different elements of the experimental block (Timm et al., 2013). Again, attenuation and attention effects on the N1 were found, but the attenuation of self-generated tones to (temporally unpredictable) external tones was similar when tones were attended to or not. Counting the number of visual stimuli with a particular colour used for cueing, further, did not affect the N1 in cued external and cued self-generation conditions but a later negative component (Harrison et al., 2021). Considering these findings and the inconsistent suggested affordance of attention to cues and observed or performed button presses in the three studies, it does not seem likely that our N1 results are explained by attention.

For the P2, attentional effects are less clear, as the negativity of the N1-attention effect could reduce the P2 and balance out attentional effects that would increase P2 amplitudes (Horváth, 2015). For instance, Saupe et al. (2013) found no attention effect on the P2. However, in some attenuation studies the latencies could indicate an overlay with a later positive component such as the P3a (e.g., Ghio et al., 2018; Seidel et al., 2021), which has been in this context interpreted as a call for attention for unpredictable and deviant tones (e.g., Knolle et al., 2013a, 2013b; Neszmélyi & Horváth, 2017). The P2 has also been related to attentional processes, such as the withdrawal of attention (Crowley & Colrain, 2004), and its attenuation interpreted as the withdrawal of attention to predicted outcomes identified as self-generated (e.g., Harrison et al., 2025). Likewise, the P2 attenuation in our studies might be related to a withdrawal of attention to tones that were identified as the outcome of any action or event, especially when the action was perceived as intentional.

3.5 Limitations and Outlook

A confound in the self-generation condition regards the performed action, which was a button press that yielded tactile and auditory feedback itself, and has been found to be performed less forcefully when a tone is elicited by it, possibly indicating action optimisation (Horváth et al., 2018; Neszmélyi & Horváth, 2017). This is one of the reasons the motor correction applied in Studies 1, 2, and 3 has been criticised (Horváth, 2015; Neszmélyi et al., 2022), although the delay that decreased temporal proximity between action and tone in the studies may have mitigated this confound. There are further minor differences in the processing of EEG data and in extracting N1 and P2 amplitude values (e.g., Ghio et al., 2021; Mifsud & Whitford, 2017; Neszmélyi & Horváth, 2021), which may slightly decrease comparability across previous studies and the three studies of the dissertation (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). For example, the low cut-off filter applied in Studies 1 and 2 was 0.1 Hz (Egan et al., 2023; Egan et al., 2025), whereas a 0.3 Hz filter was used in Study 3 (Egan et al., 2024), which may influence the N1 and P2 amplitudes (Luck, 2014). While the three studies and other studies have focused on the N1 and P2 (e.g., Klaffehn et al., 2019; Mifsud, Beesley, et al., 2016) or even solely on the N1 (e.g., Elijah et al., 2018; Timm et al., 2013), some studies have analysed additional ERP components, such as the N2 or N2b (e.g., Harrison et al., 2025; Harrison et al., 2021; Knolle et al., 2013b; Seidel et al., 2023), which could offer further insights on similarities and differences between the implemented experimental conditions. Further insights could also be gained from considering the sub-components of N1 amplitudes (e.g., SanMiguel, Todd, et al., 2013), in particular when applying principal component analysis (e.g., Tast et al., 2024).

Another potential decrease in comparability stems from the different implementations of the action-observation condition(s) across the three studies, as well as compared to previous studies. For instance, in the live observation in Study 1 (Egan et al., 2023) a mirror construction was used to reflect the button press in first-person perspective at the height of the computer screen, unlike in the live observation of Ghio et al. (2021). While it had the purpose of increasing the comparability between the experimental conditions, it is unclear what additional effects the mirror-view on one's own action might elicit and whether it could decrease sensory attenuation. However, the lack of difference in P2 attenuation between action performance and observation was also found with a more naturalistic view on one's own and the other's action (Ghio et al., 2021). Therefore, although it could be explored in future research, it is unlikely that this poses an alternative explanation for the P2 results. The video action-observation conditions in Studies 2 and 3 (Egan et al., 2024; Egan et al., 2025) further used different stimuli, a different number of frames, and total length of button-press movement. As such, the observed button-press movements in Study 2 (Egan et al., 2025) were longer than in Study 3 (Egan et al., 2024) and were rated as moving rigidly. This was due to the latency error in frame presentation. The rigid movement in Study 2 (Egan et al., 2025) as well as the short two-frame movement in Study 3 (Egan et al., 2024) could have contributed to the diminished P2 attenuation in action observation. A high-resolution, naturalistic video of button presses might yield different results, suggesting that the degree of attenuation depends on the degree of intentionality ascribed to the agent (which could increase with the increased clarity and naturalism of the observation), as opposed to other factors that differ between the action-observation settings. In future studies, measures of intentionality as well as agency attribution might be obtained from the participants and video action observation might be implemented as a video stream of a previous participant.

It should be emphasised that the studies addressing intentionality (Roselli et al., 2022; Roselli et al., 2021), which were discussed in terms of possible explanations for the lack of agency attribution effect on the P2 in Study 2 (Egan et al., 2025), measured intentional binding. Hughes et al. (2013) proposed different underlying mechanisms for binding and attenuation, relating binding to temporal control and not identity prediction. A subsequent meta-analysis showed that indeed the action shift related to temporal control, the outcome shift related to temporal predictability, while neither showed an effect of identity prediction (Tanaka et al., 2019). Dewey and Knoblich (2014) did not find a significant correlation between attenuation (i.e., auditory perceptual attenuation via PSE) and binding (see also Borhani et al., 2017). Further, it has been proposed that, as for perceptual attenuation, binding effects arise from heuristic judgements (Reddy, 2022). For instance, studies have shown that attentional processes play a large role in binding (e.g., Schwarz & Weller, 2023). Therefore, the link between (neurophysiological) attenuation and binding effects is not clear and the suggested relevance of intentionality for P2 attenuation should be regarded with caution. Future research could, on the one hand, clarify the ambiguous relationship of binding and attenuation, and on the other hand, as stated above, examine the role of intentionality for P2 attenuation via questionnaire measures and more naturalistic action-observation implementations.

Another option for the implementation of action-observation settings might be the use of VR with EEG, as mentioned previously. VR has been explored for studying neurophysiological attenuation in the somatosensory modality, which showed a general contingency effect on the P300 and replicated sensory attenuation for earlier ERP components (Giannini et al., 2025). VR can additionally be used for eye tracking (Klotzsche et al., 2023), thus enabling an incorporated control of focused observation and preventing the need for post hoc procedures to exclude segments, in which the eyes of the participant appear to be closed, as we did in Study 2 (Egan et al., 2025). VR agents to be observed might also be adjusted to resemble the experimenter, a friend of the participant, or even the participant themselves. With VR, more ecologically valid action outcomes (e.g., Ross et al., 2017) might be implemented while maintaining strict control over stimulus features. For instance, a sports game (e.g., golf, badminton) might be implemented. In this context, the auditory stimulus might either be an informative sound (i.e., the impact sound from the shuttle on the racket informs on the quality of the hit in badminton) or an uninformative one (i.e., random or fixed, or in a different sports game, in which the impact sound is less informative, such as tennis) as a manipulation of attention. This sport can then either be played in a single format (e.g., a ball machine providing the ball) or including the observation of a VR agent. This might also be useful for the further exploration of the specific research question of Study 1 (Egan et al., 2023), regarding the potential effect of culture (Cao & Gross, 2015) on the degree of attenuation in action observation. As competitiveness was identified as one factor of individualism (Shulruf et al., 2007) and our sample scored relatively low on it, effects of action observation on PSE values and the P2 amplitudes might differ between competitive (i.e., no attenuation of observed tones) and collaborative settings (i.e., attenuation of observed tones). Using VR in more complex settings could be established with one VR agent as the collaborator and another as the competitor (e.g., playing a doubles or singles tennis match), or promoting either individualistic or collectivistic social norms. Alternatively, joint action tasks (e.g., Bolt & Loehr, 2023) might be adapted, such as to include rewards for task performance that can be won by either only one person or only the pair of them. In previous joint-action tasks, generating a tones sequence with a partner (i.e., tones every 500 ms), seemed to enhance P2 amplitudes (Bolt & Loehr, 2021), particularly for other-generated tones, the more coordination was required (Bolt & Loehr, 2023). They interpreted this as increased temporal orienting of attention to the tones. On the other hand, a study by Weiss et al. (2011b), in which one cued the other's action, showed reduced PSE values in the interactive setting, which was interpreted as an increased feeling of agency to contrast with the agency of the other in determining the temporal onset (see also Neszmélyi & Horváth, 2021, for an action-cueing task in an EEG study). Therefore, interaction might introduce further effects of attention and agency that need to be considered when designing more complex action-observation settings in VR.

Lastly, the focus in this dissertation was on the auditory modality, specifically on the processing of tones, but some studies have examined attenuation in action observation in other modalities raising further research questions. For instance, Kilteni et al. (2021) did not find somatosensory attenuation in action observation in a series of behavioural experiments (e.g., via PSE values) but Kilteni et al. (2018) did find attenuation in a motor-imagery condition in a force-matching task. How measures of attenuation in different modalities relate to each other (e.g., Schafer & Marcus, 1973) and how motor imagery, overlapping in activation with action observation and performance (Hardwick et al., 2018), compares to action observation (e.g., Tang et al., 2025) in its resulting attenuation could be explored in future studies.

3.6 Conclusion

Sensory attenuation has been interpreted as the result of predictive mechanisms, in that the match of the perceived sensory input and the sensory input predicted based on the performed action reduces sensory processing (e.g., Korka et al., 2022). Because of the overlap of activation during action observation and performance (e.g., Hardwick et al., 2018), explanatory models of attenuation proposed the same resource for predicting sensory outcomes in action performance and observation (Pickering & Clark, 2014). Neurophysiological attenuation of tones generated by an observed person compared to external tones was partially found in previous studies (Ghio et al., 2021; Ghio et al., 2018). The aim of the three studies of this dissertation was to gain further insights from comparing N1 and P2 attenuation of self-generated tones and of tones generated by an observed action, while accounting for temporal predictability.

Individualism was assessed in Study 1 (Egan et al., 2023) as a potentially influencing factor of neurophysiological attenuation in action observation. Individualism did not significantly affect N1 or P2 attenuation in action observation, although this was the case in a previous study for perceptual attenuation measured via the PSE (Cao & Gross, 2015). Future studies could further test the effect of individualism in samples with more diverging individualism and collectivism scores, and examine whether different mechanisms underly neurophysiological and perceptual attenuation (e.g., Ody et al., 2023).

Action-observation conditions with different agents were implemented in Study 2 (Egan et al., 2025) to investigate the role of agency attribution for P2 attenuation. However, we found no difference in P2 attenuation across the cue and action-observation conditions. P2 amplitudes in the (human) action-observation condition differed from P2 amplitudes in the cue condition only in Study 1 (Egan et al., 2023), which implemented a live action-observation setting in first-person perspective as opposed to the video action-observation setting in first-person perspective in Studies 2 and 3 (Egan et al., 2024; Egan et al., 2025). Potential explanations regard the possibly stronger vicarious sense of agency and attribution of intentionality in the live setting. Future research could examine these potential effects, for instance by utilising VR, and thus contribute to understanding the functional significance of the P2.

In Study 3 (Egan et al., 2024), whether-prediction errors were induced by reducing the contingency of action and outcome to examine the effects of outcome uncertainty on attenuation in action performance and action observation. Whereas the overall attenuation with 50% contingency showed a diverging pattern, a similar effect of the trial-by-trial varying prediction error was found for N1 and P2 amplitudes. Smaller whether-prediction errors yielded smaller amplitudes, and this effect was stronger for self-generation. This might be due to an increased salience of own action outcomes (e.g., Knolle et al., 2013b) and a more rewarding experience when predictions of sensory consequences of own actions prove correct (see Ripollés et al., 2014) or simply when enacting control over the environment by producing action effects (e.g., Eitam et al., 2013). With the 50% contingency, the P2 amplitudes of self-generated tones were enhanced compared to cued and observed action outcomes (i.e., as large as the P2 of external tones). Such (self-specific) contingency effects on P2 amplitudes were previously found (Han et al., 2022; Seidel et al., 2021). Unlike the P2, N1 amplitudes were attenuated in action performance and observation compared to external tones, which might indicate the involvement of unspecific mechanisms, such as sensorimotor gating (cf. Tast et al., 2025).

Importantly, N1 amplitudes were not consistently attenuated compared to external tones, as they were attenuated only for cued external tones in Study 1 (Egan et al., 2023), only for tones generated by an observed human agent in Study 2 (Egan et al., 2025), and only for self-generated and tones generated by an observed person in Study 3 (Egan et al., 2024), even though each study included self-generated and cued tones as well as tones generated by an observed person. Yet, N1 amplitudes in these conditions did not differ significantly. Thus, our results tentatively supported previous studies linking N1 attenuation to temporal predictability (e.g., Harrison et al., 2021) but also suggested different underlying mechanisms for self-generation (i.e., stronger effect of whether-prediction errors), despite yielding similar attenuation as action observation (Egan et al., 2024). This is in line with studies showing, particularly for the N1, the involvement of a cerebellar forward model for self-generation (e.g., Cao et al., 2017; Ghio et al., 2025; Knolle et al., 2013a) but not action observation (Ghio et al., 2025).

For the P2, we found an attenuating effect of temporal predictability in all three studies (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). However, self-generated tones showed specific P2 effects, in that there was stronger attenuation in Studies 1 and 2 (Egan et al., 2023; Egan et al., 2025) compared to cued tones and also compared to action observation in Study 2 (Egan et al., 2025), and no attenuation when there was a 50% contingency between action and tone (Egan et al., 2024). These results might be linked to the interpretation of the P2 attenuation as a more retrospective agency judgement (e.g., Knolle et al., 2013a; Seidel et al., 2021; Timm et al., 2016) and the degree of intentionality attributed to the agent, the contingency between action and outcome as well as (consistent) temporal proximity could be relevant to the identification of the tone as an outcome of an event, particularly of own action.

In sum, auditory outcomes of observed actions seem to be processed not like auditory outcomes of own actions but like auditory outcomes of other visual cues, although this need not be reflected in the degree of (N1) attenuation and (for the P2) might depend on the circumstances of the action observation (Egan et al., 2023; Egan et al., 2024; Egan et al., 2025). It should be kept in mind that observed similarities in amplitude do not necessarily imply equal underlying processes (Press et al., 2023) and multiple mechanisms might lead to attenuation, including specific and unspecific ones (cf. Tast et al., 2025) as well as multiple neural networks (Gu et al., 2025). Our results are in line with accounts of neurophysiological attenuation as a combination of different predictive mechanisms, including motor-based and general prediction (e.g., Korka et al., 2022) and call for further research on sensory outcome processing in action observation.

4 References

Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience*, *21*(4), 791-802. https://doi.org/10.1162/jocn.2009.21055

Angelini, M., Del Vecchio, M., Lopomo, N. F., Gobbo, M., & Avanzini, P. (2021). Perspective-dependent activation of frontoparietal circuits during the observation of a static body effector. *Brain Research*, *1769*, 147604. https://doi.org/10.1016/j.brainres.2021.147604

Angelini, M., Fabbri-Destro, M., Lopomo, N. F., Gobbo, M., Rizzolatti, G., & Avanzini, P. (2018). Perspective-dependent reactivity of sensorimotor mu rhythm in alpha and beta ranges during action observation: an EEG study. *Scientific Reports*, *8*(1), 12429. https://doi.org/10.1038/s41598-018-30912-w

Antonioni, A., Raho, E. M., Straudi, S., Granieri, E., Koch, G., & Fadiga, L. (2024). The cerebellum and the Mirror Neuron System: A matter of inhibition? From neurophysiological evidence to neuromodulatory implications. A narrative review. *Neuroscience and Biobehavioral Reviews*, *164*, 105830. https://doi.org/10.1016/j.neubiorev.2024.105830

Bäß, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*(2), 137-143. https://doi.org/10.1016/j.ijpsycho.2008.06.005

Behroozmand, R., Sangtian, S., Korzyukov, O., & Larson, C. R. (2016). A temporal predictive code for voice motor control: Evidence from ERP and behavioral responses to pitch-shifted auditory feedback. *Brain Research*, *1636*, 1-12. https://doi.org/10.1016/j.brainres.2016.01.040

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). The Cardiff Anomalous Perceptions Scale (CAPS): A New Validated Measure of Anomalous Perceptual Experience. *Schizophrenia Bulletin*, *32*(2), 366-377. https://doi.org/10.1093/schbul/sbj014

Bellebaum, C., Jokisch, D., Gizewski, E. R., Forsting, M., & Daum, I. (2012). The neural coding of expected and unexpected monetary performance outcomes: dissociations between active and observational learning. *Behavioural Brain Research*, *227*(1), 241-251. https://doi.org/10.1016/j.bbr.2011.10.042

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: event-related brain potentials in active and observational learning from feedback. *Cerebral Cortex*, *20*(12), 2874-2883. https://doi.org/10.1093/cercor/bhq038

Blakemore, S.-J., Smith, J., Steel, R., Johnstone, C. E., & Frith, C. D. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychological Medicine*, *30*(5), 1131-1139. https://doi.org/10.1017/s0033291799002676

Blakemore, S.-J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *NeuroReport*, *11*(11), R11-R16. https://doi.org/10.1097/00001756-200008030-00002

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635-640. https://doi.org/10.1038/2870

Bolt, N. K., & Loehr, J. D. (2021). Sensory Attenuation of the Auditory P2 Differentiates Self- from Partner-Produced Sounds during Joint Action. *Journal of Cognitive Neuroscience*, *33*(11), 2297-2310. https://doi.org/10.1162/jocn_a_01760

Bolt, N. K., & Loehr, J. D. (2023). The auditory P2 differentiates self- from partner-produced sounds during joint action: Contributions of self-specific attenuation and temporal orienting of attention. *Neuropsychologia*, *182*, 108526. https://doi.org/10.1016/j.neuropsychologia.2023.108526

Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, Doing, and Controlling: Implicit Sense of Agency Over Somatosensory Events. *Psychological Science*, *28*(7), 882-893. https://doi.org/10.1177/0956797617697693

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, *14*(4), 411-427. https://doi.org/10.1007/s10339-013-0571-3

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, *13*(2), 400-404. https://doi.org/10.1111/j.1460-9568.2001.01385.x

Calvo-Merino, B., Grezes, 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*(19), 1905-1910. https://doi.org/10.1016/j.cub.2006.07.065

Cao, L., & Gross, J. (2015). Cultural Differences in Perceiving Sounds Generated by Others: Self Matters. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01865

Cao, L., Veniero, D., Thut, G., & Gross, J. (2017). Role of the Cerebellum in Adaptation to Delayed Action Effects. *Current Biology*, *27*(16), 2442-2451 e2443. https://doi.org/10.1016/j.cub.2017.06.074

Caruana, N., & McArthur, G. (2019). The mind minds minds: The effect of intentional stance on the neural encoding of joint attention. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(6), 1479-1491. https://doi.org/10.3758/s13415-019-00734-y

Casiraghi, L., Alahmadi, A. A. S., Monteverdi, A., Palesi, F., Castellazzi, G., Savini, G., Friston, K., Gandini Wheeler-Kingshott, C. A. M., & D'Angelo, E. (2019). I See Your Effort: Force-Related BOLD Effects in an Extended Action Execution-Observation Network Involving the Cerebellum. *Cerebral Cortex*, *29*(3), 1351-1368. https://doi.org/10.1093/cercor/bhy322

Charpentier, C. J., & O'Doherty, J. P. (2018). The application of computational models to social neuroscience: promises and pitfalls. *Social Neuroscience*, *13*(6), 637-647. https://doi.org/10.1080/17470919.2018.1518834

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204. https://doi.org/10.1017/S0140525X12000477

Corradini, A., & Antonietti, A. (2013). Mirror neurons and their function in cognitively understood empathy. *Consciousness and Cognition*, *22*(3), 1152-1161. https://doi.org/10.1016/j.concog.2013.03.003

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732-744. https://doi.org/10.1016/j.clinph.2003.11.021

Curio, G., Neuloh, G., Numminen, J., Jousmäki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*, *9*(4), 183-191. https://doi.org/10.1002/(SICI)1097-0193(200004)9:4%3C183::AID-HBM1%3E3.0.CO;2-Z

Dennett, D. C. (1971). Intentional Systems. *The Journal of Philosophy*, *68*(4), 87-106. https://doi.org/10.2307/2025382

Desantis, A., Weiss, C., Schutz-Bosbach, S., & Waszak, F. (2012). Believing and perceiving: authorship belief modulates sensory attenuation. *PLOS ONE*, *7*(5), e37959. https://doi.org/10.1371/journal.pone.0037959

Dewey, J. A., & Knoblich, G. (2014). Do implicit and explicit measures of the sense of agency measure the same thing? *PLOS ONE*, *9*(10), e110118. https://doi.org/10.1371/journal.pone.0110118

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*, 176-180. https://doi.org/10.1007/BF00230027

Dogge, M., Custers, R., & Aarts, H. (2019). Moving forward: On the Limits of Motor-Based Forward Models. *Trends in Cognitive Sciences*, *23*(9), 743-753. https://doi.org/10.1016/j.tics.2019.06.008

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. https://doi.org/10.1016/j.neuropsychologia.2018.12.007

Egan, S., Ghio, M., & Bellebaum, C. (2023). Auditory N1 and P2 attenuation in action observation: An event-related potential study considering effects of temporal predictability and individualism. *Biological Psychology*, *180*, 108575. https://doi.org/10.1016/j.biopsycho.2023.108575

Egan, S., Seidel, A., Weber, C., Ghio, M., & Bellebaum, C. (2024). Fifty Percent of the Time, Tones Come Every Time: Stronger Prediction Error Effects on Neurophysiological Sensory Attenuation for Self-generated Tones. *Journal of Cognitive Neuroscience*, *36*(10), 2067-2083. https://doi.org/10.1162/jocn_a_02226

Egan, S., Weber, C., Ghio, M., & Bellebaum, C. (2025). I, You, Robot: Attenuation for Auditory Outcomes of Actions Performed by Different Agents Shows Distinct Patterns for N1 and P2 Amplitudes [Manuscript submitted for publication].

Eitam, B., Kennedy, P. M., & Tory Higgins, E. (2013). Motivation from control. *Experimental Brain Research*, *229*(3), 475-484. https://doi.org/10.1007/s00221-012-3370-7

Elijah, R. B., Le Pelley, M. E., & Whitford, T. J. (2016). Modifying temporal expectations: Changing cortical responsivity to delayed self-initiated sensations with training. *Biological Psychology*, *120*, 88-95. https://doi.org/10.1016/j.biopsycho.2016.09.001

Elijah, R. B., Le Pelley, M. E., & Whitford, T. J. (2018). Act Now, Play Later: Temporal Expectations Regarding the Onset of Self-initiated Sensations Can Be Modified with Behavioral Training. *Journal of Cognitive Neuroscience*, *30*(8), 1145-1156. https://doi.org/10.1162/jocn_a_01269

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864-886. https://doi.org/10.1037/0033-295X.114.4.864

Feder, S., Miksch, J., Grimm, S., Krems, J. F., & Bendixen, A. (2023). Using event-related brain potentials to evaluate motor-auditory latencies in virtual reality. *Frontiers in Neuroergonomics*, *4*, 1196507. https://doi.org/10.3389/fnrgo.2023.1196507

Feinberg, I. (1978). Efference copy and corollary discharge: implications for thinking and its disorders. *Schizophrenia Bulletin*, *4*(4), 636-640. https://doi.org/10.1093/schbul/4.4.636

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., & Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology*, *44*(4), 522-529. https://doi.org/10.1111/j.1469-8986.2007.00533.x

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological Evidence of Corollary Discharge Dysfunction in

Schizophrenia. *American Journal of Psychiatry*, *158*(12), 2069-2071. https://doi.org/10.1176/appi.ajp.158.12.2069

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*(4), 804-812. https://doi.org/10.1093/schbul/sbt072

Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Daléry, J., d'Amato, T., & Jeannerod, M. (2001). Defective Recognition of One's Own Actions in Patients With Schizophrenia. *American Journal of Psychiatry*, *158*(3), 454-459. https://doi.org/10.1176/appi.ajp.158.3.454

Fritz, C., Flick, M., & Zimmermann, E. (2022). Tactile motor attention induces sensory attenuation for sounds. *Consciousness and Cognition*, *104*, 103386. https://doi.org/10.1016/j.concog.2022.103386

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14-21. https://doi.org/10.1016/S1364-6613(99)01417-5

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593-609. https://doi.org/10.1093/brain/119.2.593

Ge, S., Liu, H., Lin, P., Gao, J., Xiao, C., & Li, Z. (2018). Neural Basis of Action Observation and Understanding From First- and Third-Person Perspectives: An fMRI Study. *Frontiers in Behavioral Neuroscience*, *12*, 283. https://doi.org/10.3389/fnbeh.2018.00283

Gentsch, A., Schütz-Bosbach, S., Endrass, T., & Kathmann, N. (2012). Dysfunctional forward model mechanisms and aberrant sense of agency in obsessive-compulsive disorder. *Biological Psychiatry*, *71*(7), 652-659. https://doi.org/10.1016/j.biopsych.2011.12.022

Ghio, M., Egan, S., & Bellebaum, C. (2021). Similarities and differences between performers and observers in processing auditory action consequences: Evidence from simultaneous EEG acquisition. *Journal of Cognitive Neuroscience*, *33*(4), 683-694. https://doi.org/10.1162/jocn_a_01671

Ghio, M., Haegert, K., Seidel, A., Suchan, B., Thoma, P., & Bellebaum, C. (2025). The prediction of auditory consequences of own and observed actions: a brain decoding multivariate pattern study. *Cerebral Cortex*, *35*(4). https://doi.org/10.1093/cercor/bhaf091

Ghio, M., Scharmach, K., & Bellebaum, C. (2018). ERP correlates of processing the auditory consequences of own versus observed actions. *Psychophysiology*, *55*(6), e13048. https://doi.org/10.1111/psyp.13048

Giannini, G., Nierhaus, T., & Blankenburg, F. (2025). Investigation of sensory attenuation in the somatosensory domain using EEG in a novel virtual reality paradigm. *Scientific Reports*, *15*(1), 2819. https://doi.org/10.1038/s41598-025-87244-9

Gu, J., Buidze, T., Zhao, K., Glascher, J., & Fu, X. (2025). The neural network of sensory attenuation: A neuroimaging meta-analysis. *Psychonomic Bulletin & Review*, *32*(1), 31-51. https://doi.org/10.3758/s13423-024-02532-1

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382-385. https://doi.org/10.1038/nn827

Han, N., Jack, B. N., Hughes, G., Elijah, R. B., & Whitford, T. J. (2021). Sensory attenuation in the absence of movement: Differentiating motor action from sense of agency. *Cortex*, *141*, 436-448. https://doi.org/10.1016/j.cortex.2021.04.010

Han, N., Jack, B. N., Hughes, G., & Whitford, T. J. (2022). The Role of Action-Effect Contingency on Sensory Attenuation in the Absence of Movement. *Journal of Cognitive Neuroscience*, *34*(8), 1488-1499. https://doi.org/10.1162/jocn_a_01867

Hardwick, R. M., Caspers, S., Eickhoff, S. B., & Swinnen, S. P. (2018). Neural correlates of action: Comparing meta-analyses of imagery, observation, and execution. *Neuroscience and Biobehavioral Reviews*, *94*, 31-44. https://doi.org/10.1016/j.neubiorev.2018.08.003

Harrison, A. W., Christensen, B. K., & Whitford, T. J. (2025). Misprediction Sensitivity Is Diminished for Action-effects Resulting from Stimulus-driven Movement. *Journal of Cognitive Neuroscience*, *37*(9), 1431-1448. https://doi.org/10.1162/jocn_a_02321

Harrison, A. W., Hughes, G., Rudman, G., Christensen, B. K., & Whitford, T. J. (2023). Exploring the internal forward model: action-effect prediction and attention in sensorimotor processing. *Cerebral Cortex*, *33*(14), 9130-9143. https://doi.org/10.1093/cercor/bhad189

Harrison, A. W., Mannion, D. J., Jack, B. N., Griffiths, O., Hughes, G., & Whitford, T. J. (2021). Sensory attenuation is modulated by the contrasting effects of predictability and control. *NeuroImage*, *237*, 118103. https://doi.org/10.1016/j.neuroimage.2021.118103

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of Imprecise Corollary Discharge in Schizophrenia to Auditory Hallucinations. *Archives of General Psychiatry*, *64*(3), 286-296. https://doi.org/10.1001/archpsyc.64.3.286

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture*, *2*(1). https://doi.org/10.9707/2307-0919.1014

Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54-65. https://doi.org/10.1016/j.brainres.2015.03.038

Horváth, J., Bíró, B., & Neszmélyi, B. (2018). Action-effect related motor adaptation in interactions with everyday devices. *Scientific Reports*, *8*(1), 6592. https://doi.org/10.1038/s41598-018-25161-w

Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action–sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience*, *24*(9), 1919-1931. https://doi.org/10.1162/jocn_a_00215

Hughes, G., Desantis, A., & Waszak, F. (2013). 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. https://doi.org/10.1037/a0028566

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, *60*, 653-670. https://doi.org/10.1146/annurev.psych.60.110707.163604

Jack, B. N., Chilver, M. R., Vickery, R. M., Birznieks, I., Krstanoska-Blazeska, K., Whitford, T. J., & Griffiths, O. (2021). Movement Planning Determines Sensory Suppression: An Event-related Potential Study. *Journal of Cognitive Neuroscience*, *33*(12), 2427-2439. https://doi.org/10.1162/jocn_a_01747

Kaiser, J., & Schütz-Bosbach, S. (2018). Sensory attenuation of self-produced signals does not rely on self-specific motor predictions. *European Journal of Neuroscience*, *47*(11), 1303-1310. https://doi.org/10.1111/ejn.13931

Kiepe, F., Kraus, N., & Hesselmann, G. (2021). Sensory Attenuation in the Auditory Modality as a Window Into Predictive Processing. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.704668

Kiepe, F., Kraus, N., & Hesselmann, G. (2024). Self-initiation enhances perceptual processing of auditory stimuli in an online study. *Attention, Perception, & Psychophysics*, *86*(2), 587-601. https://doi.org/10.3758/s13414-023-02827-w

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159-166. https://doi.org/10.1007/s10339-007-0170-2

Kilteni, K., Andersson, B. J., Houborg, C., & Ehrsson, H. H. (2018). Motor imagery involves predicting the sensory consequences of the imagined movement. *Nature Communications*, *9*(1), 1617. https://doi.org/10.1038/s41467-018-03989-0

Kilteni, K., & Ehrsson, H. H. (2020). Functional Connectivity between the Cerebellum and Somatosensory Areas Implements the Attenuation of Self-Generated Touch. *Journal of Neuroscience*, *40*(4), 894-906. https://doi.org/10.1523/JNEUROSCI.1732-19.2019

Kilteni, K., Engeler, P., Boberg, I., Maurex, L., & Ehrsson, H. H. (2021). No evidence for somatosensory attenuation during action observation of self-touch. *European Journal of Neuroscience*, *54*(7), 6422-6444. https://doi.org/10.1111/ejn.15436

Kilteni, K., Engeler, P., & Ehrsson, H. H. (2020). Efference Copy Is Necessary for the Attenuation of Self-Generated Touch. *iScience*, *23*(2), 100843. https://doi.org/10.1016/j.isci.2020.100843

Klaffehn, A. L., Baess, P., Kunde, W., & Pfister, R. (2019). Sensory attenuation prevails when controlling for temporal predictability of self-and externally generated tones. *Neuropsychologia*, *132*, 107145. https://doi.org/10.1016/j.neuropsychologia.2019.107145

Klotzsche, F., Gaebler, M., Villringer, A., Sommer, W., Nikulin, V., & Ohl, S. (2023). Visual short-term memory-related EEG components in a virtual reality setup. *Psychophysiology*, *60*(11), e14378. https://doi.org/10.1111/psyp.14378

Knolle, F., Schröger, E., & Kotz, S. A. (2013a). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*(9), 2449-2461. https://doi.org/10.1016/j.cortex.2012.12.012

Knolle, F., Schröger, E., & Kotz, S. A. (2013b). Prediction errors in self-and externally-generated deviants. *Biological Psychology*, *92*(2), 410-416. https://doi.org/10.1016/j.biopsycho.2012.11.017

Kobza, S., & Bellebaum, C. (2015). Processing of action- but not stimulus-related prediction errors differs between active and observational feedback learning. *Neuropsychologia*, *66*, 75-87. https://doi.org/10.1016/j.neuropsychologia.2014.10.036

Korka, B., Widmann, A., Waszak, F., Darriba, Á., & Schröger, E. (2022). The auditory brain in action: Intention determines predictive processing in the auditory system—A review of current paradigms and findings. *Psychonomic Bulletin & Review*, *29*(2), 321-342. https://doi.org/10.3758/s13423-021-01992-z

Kort, N. S., Ford, J. M., Roach, B. J., Gunduz-Bruce, H., Krystal, J. H., Jaeger, J., Reinhart, R. M., & Mathalon, D. H. (2017). Role of N-Methyl-D-Aspartate Receptors in Action-Based Predictive Coding Deficits in Schizophrenia. *Biological Psychiatry*, *81*(6), 514-524. https://doi.org/10.1016/j.biopsych.2016.06.019

Koverola, M., Kunnari, A., Sundvall, J., & Laakasuo, M. (2022). General Attitudes Towards Robots Scale (GAToRS): A New Instrument for Social Surveys. *International Journal of Social Robotics*, *14*(7), 1559-1581. https://doi.org/10.1007/s12369-022-00880-3

Kühn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., & Voss, M. (2011). Whodunnit? Electrophysiological correlates of agency judgements. *PLOS ONE*, *6*(12), e28657. https://doi.org/10.1371/journal.pone.0028657

Lamm, C., & Majdandzic, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy - A critical comment. *Neuroscience Research*, *90*, 15-24. https://doi.org/10.1016/j.neures.2014.10.008

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, *69*(1), 127-137. https://doi.org/10.1016/j.bandc.2008.06.004

Lange, K. (2011). The reduced N1 to self-generated tones: an effect of temporal predictability? *Psychophysiology*, *48*(8), 1088-1095. https://doi.org/10.1111/j.1469-8986.2010.01174.x

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the Empathy Quotient. *Psychological Medicine*, *34*(5), 911-920. https://doi.org/10.1017/S0033291703001624

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Gaspelin, N. (2017). How to Get Statistically Significant Effects in Any ERP Experiment (and Why You Shouldn't). *Psychophysiology*, *54*(1), 146-157. https://doi.org/10.1111/psyp.12639

Mariano, M., Rossetti, I., Maravita, A., Paulesu, E., & Zapparoli, L. (2024). Sensory Attenuation Deficit and Auditory Hallucinations in Schizophrenia: A Causal Mechanism or a Risk Factor? Evidence From Meta-Analyses on the N1 Event-Related Potential Component. *Biological Psychiatry*, *96*(3), 207-221. https://doi.org/10.1016/j.biopsych.2023.12.026

McCarthy, G., & Donchin, E. (1976). The effects of temporal and event uncertainty in determining the waveforms of the auditory event related potential (ERP). *Psychophysiology*, *13*(6), 581-590. https://doi.org/10.1111/j.1469-8986.1976.tb00885.x

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, *14*(17), 2135-2137. https://doi.org/10.1097/00001756-200312020-00001

Mifsud, N. G., Beesley, T., Watson, T. L., Elijah, R. B., Sharp, T. S., & Whitford, T. J. (2018). Attenuation of visual evoked responses to hand and saccade-initiated flashes. *Cognition*, *179*, 14-22. https://doi.org/10.1016/j.cognition.2018.06.005

Mifsud, N. G., Beesley, T., Watson, T. L., & Whitford, T. J. (2016). Attenuation of auditory evoked potentials for hand and eye-initiated sounds. *Biological Psychology*, *120*, 61-68. https://doi.org/10.1016/j.biopsycho.2016.08.011

Mifsud, N. G., Oestreich, L. K. L., Jack, B. N., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2016). Self-initiated actions result in suppressed auditory but amplified visual evoked components in healthy participants. *Psychophysiology*, *53*(5), 723-732. https://doi.org/10.1111/psyp.12605

Mifsud, N. G., & Whitford, T. J. (2017). Sensory attenuation of self-initiated sounds maps onto habitual associations between motor action and sound. *Neuropsychologia*, *103*, 38-43. https://doi.org/10.1016/j.neuropsychologia.2017.07.019

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, *17*(1), 136-144. https://doi.org/10.1016/j.concog.2006.12.004

Moore, J. W., Dickinson, A., & Fletcher, P. C. (2011). Sense of agency, associative learning, and schizotypy. *Consciousness and Cognition*, *20*(3), 792-800. https://doi.org/10.1016/j.concog.2011.01.002

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and Cognition*, *21*(1), 546-561. https://doi.org/10.1016/j.concog.2011.12.002

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, *20*(8), 750-756. https://doi.org/10.1016/j.cub.2010.02.045

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure.

Psychophysiology, *24*(4), 375-425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x

Neave, N., Jackson, R., Saxton, T., & Hönekopp, J. (2015). The influence of anthropomorphic tendencies on human hoarding behaviours. *Personality and Individual Differences*, *72*, 214-219. https://doi.org/10.1016/j.paid.2014.08.041

Neszmélyi, B., & Horváth, J. (2017). Consequences matter: Self-induced tones are used as feedback to optimize tone-eliciting actions. *Psychophysiology*, *54*(6), 904-915. https://doi.org/10.1111/psyp.12845

Neszmélyi, B., & Horváth, J. (2021). Processing and utilization of auditory action effects in individual and social tasks. *Acta Psychologica*, *217*, 103326. https://doi.org/10.1016/j.actpsy.2021.103326

Neszmélyi, B., Nevelo, D., & Horváth, J. (2022). Turning a blind eye to motor differences leads to bias in estimating action-related auditory ERP attenuation. *Biological Psychology*, *173*, 108387. https://doi.org/10.1016/j.biopsycho.2022.108387

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, *97*(2), 913-918. https://doi.org/10.1073/pnas.97.2.913

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science*, *304*(5669), 452-454. https://doi.org/10.1126/science.1094285

Ody, E., Straube, B., He, Y., & Kircher, T. (2023). Perception of self-generated and externally-generated visual stimuli: Evidence from EEG and behavior. *Psychophysiology*, *60*(8), e14295. https://doi.org/10.1111/psyp.14295

Oestreich, L. K., Mifsud, N. G., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2015). Subnormal sensory attenuation to self-generated speech in schizotypy: Electrophysiological evidence for a 'continuum of psychosis'. *International Journal of Psychophysiology*, *97*(2), 131-138. https://doi.org/10.1016/j.ijpsycho.2015.05.014

Orpella, J., Mas-Herrero, E., Ripollés, P., Marco-Pallarés, J., & de Diego-Balaguer, R. (2021). Language statistical learning responds to reinforcement learning principles rooted in the striatum. *PLOS Biology*, *19*(9), e3001119. https://doi.org/10.1371/journal.pbio.3001119

Palmer, C. E., Davare, M., & Kilner, J. M. (2016). Physiological and Perceptual Sensory Attenuation Have Different Underlying Neurophysiological Correlates. *Journal of Neuroscience*, *36*(42), 10803-10812. https://doi.org/10.1523/JNEUROSCI.1694-16.2016

Paraskevoudi, N., & SanMiguel, I. (2021). Self-generation and sound intensity interactively modulate perceptual bias, but not perceptual sensitivity. *Scientific Reports*, *11*(1), 17103. https://doi.org/10.1038/s41598-021-96346-z

Penton, T., Wang, X., Coll, M. P., Catmur, C., & Bird, G. (2018). The influence of action-outcome contingency on motivation from control. *Experimental Brain Research*, *236*(12), 3239-3249. https://doi.org/10.1007/s00221-018-5374-4

Pereira, D. R., Cardoso, S., Ferreira-Santos, F., Fernandes, C., Cunha-Reis, C., Paiva, T. O., Almeida, P. R., Silveira, C., Barbosa, F., & Marques-Teixeira, J. (2014). Effects of inter-stimulus interval (ISI) duration on the N1 and P2 components of the auditory event-related potential. *International Journal of Psychophysiology*, *94*(3), 311-318. https://doi.org/10.1016/j.ijpsycho.2014.09.012

Peters, E., Joseph, S., Day, S., & Garety, P. (2004). Measuring Delusional Ideation: The 21-Item Peters et al. Delusions Inventory (PDI). *Schizophrenia Bulletin*, *30*(4), 1005-1022. https://doi.org/10.1093/oxfordjournals.schbul.a007116

Peters, E. R., Joseph, S. A., & Garety, P. A. (1999). Measurement of delusional ideation in the normal population: introducing the PDI (Peters et al. Delusions

Inventory). *Schizophrenia Bulletin*, *25*(3), 553-576. https://doi.org/10.1093/oxfordjournals.schbul.a033401

Pickering, M. J., & Clark, A. (2014). Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, *18*(9), 451-456. https://doi.org/10.1016/j.tics.2014.05.006

Pinheiro, A. P., Schwartze, M., Gutierrez, F., & Kotz, S. A. (2019). When temporal prediction errs: ERP responses to delayed action-feedback onset. *Neuropsychologia*, *134*, 107200. https://doi.org/10.1016/j.neuropsychologia.2019.107200

Poonian, S. K., & Cunnington, R. (2013). Intentional binding in self-made and observed actions. *Experimental Brain Research*, *229*(3), 419-427. https://doi.org/10.1007/s00221-013-3505-5

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*(4), 752-764. https://doi.org/10.1162/jocn_a_00745

Press, C., Kok, P., & Yon, D. (2020). The Perceptual Prediction Paradox. *Trends in Cognitive Sciences*, *24*(1), 13-24. https://doi.org/10.1016/j.tics.2019.11.003

Press, C., Thomas, E. R., & Yon, D. (2023). Cancelling cancellation? Sensorimotor control, agency, and prediction. *Neuroscience and Biobehavioral Reviews*, *145*, 105012. https://doi.org/10.1016/j.neubiorev.2022.105012

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. In R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Raine, A. (1991). The SPQ: A Scale for the Assessment of Schizotypal Personality Based on DSM-III-R Criteria. *Schizophrenia Bulletin*, *17*(4), 555-564. https://doi.org/10.1093/schbul/17.4.555

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79-87. https://doi.org/10.1038/4580

Reddy, N. N. (2022). The implicit sense of agency is not a perceptual effect but is a judgment effect. *Cognitive Processing*, *23*(1), 1-13. https://doi.org/10.1007/s10339-021-01066-x

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasky (Eds.), *Classical conditioning II: Current Research and Theory* (pp. 64-99). Appleton-Century-Crofts.

Reznik, D., Henkin, Y., Levy, O., & Mukamel, R. (2015). Perceived loudness of self-generated sounds is differentially modified by expected sound intensity. *PLOS ONE*, *10*(5), e0127651. https://doi.org/10.1371/journal.pone.0127651

Reznik, D., Henkin, Y., Schadel, N., & Mukamel, R. (2014). Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nature Communications*, *5*(1), 4059. https://doi.org/10.1038/ncomms5059

Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H.-J., Rodríguez-Fornells, A., & Noesselt, T. (2014). The role of reward in word learning and its implications for language acquisition. *Current Biology*, *24*(21), 2606-2611. https://doi.org/10.1016/j.cub.2014.09.044

Rizzolatti, G., & Fogassi, L. (2014). The mirror mechanism: recent findings and perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130420. https://doi.org/10.1098/rstb.2013.0420

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*(4), 264-274. https://doi.org/10.1038/nrn2805

Roselli, C., Ciardo, F., De Tommaso, D., & Wykowska, A. (2022). Human-likeness and attribution of intentionality predict vicarious sense of agency over humanoid robot actions. *Scientific Reports*, *12*(1), 13845. https://doi.org/10.1038/s41598-022-18151-6

Roselli, C., Ciardo, F., & Wykowska, A. (2021). Intentions with actions: The role of intentionality attribution on the vicarious sense of agency in Human-Robot interaction. *Quarterly Journal of Experimental Psychology*, *75*(4), 616-632. https://doi.org/10.1177/17470218211042003

Ross, B., Barat, M., & Fujioka, T. (2017). Sound-Making Actions Lead to Immediate Plastic Changes of Neuromagnetic Evoked Responses and Induced beta-Band Oscillations during Perception. *Journal of Neuroscience*, *37*(24), 5948-5959. https://doi.org/10.1523/JNEUROSCI.3613-16.2017

Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. *Neuropsychologia*, *51*(5), 922-929. https://doi.org/10.1016/j.neuropsychologia.2013.02.005

SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N 1 component of the auditory ERP. *Psychophysiology*, *50*(4), 334-343. https://doi.org/10.1111/psyp.12024

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schroger, E. (2013). Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. *Journal of Neuroscience*, *33*(20), 8633-8639. https://doi.org/10.1523/JNEUROSCI.5821-12.2013

Sato, A. (2008). Action observation modulates auditory perception of the consequence of others' actions. *Consciousness and Cognition*, *17*(4), 1219-1227. https://doi.org/10.1016/j.concog.2008.01.003

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, *94*(3), 241-255. https://doi.org/10.1016/j.cognition.2004.04.003

Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial suppression of self-generated sounds and its dependence on attention. *International Journal of Psychophysiology*, *90*(3), 300-310. https://doi.org/10.1016/j.ijpsycho.2013.09.006

Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain Responses. *Science*, *181*(4095), 175-177. https://doi.org/10.1126/science.181.4095.175

Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: a lesson from cognitive psychophysiology. *European Journal of Neuroscience*, *41*(5), 641-664. https://doi.org/10.1111/ejn.12816

Schwarz, K. A., & Weller, L. (2023). Distracted to a fault: Attention, actions, and time perception. *Attention, Perception, & Psychophysics*, *85*(2), 301-314. https://doi.org/10.3758/s13414-022-02632-x

Sebastiani, V., de Pasquale, F., Costantini, M., Mantini, D., Pizzella, V., Romani, G. L., & Della Penna, S. (2014). Being an agent or an observer: different spectral dynamics revealed by MEG. *NeuroImage*, *102 Pt 2*, 717-728. https://doi.org/10.1016/j.neuroimage.2014.08.031

Seidel, A., Ghio, M., Studer, B., & Bellebaum, C. (2021). Illusion of control affects ERP amplitude reductions for auditory outcomes of self-generated actions. *Psychophysiology*, *58*(5), e13792. https://doi.org/10.1111/psyp.13792

Seidel, A., Weber, C., Ghio, M., & Bellebaum, C. (2023). My view on your actions: Dynamic changes in viewpoint-dependent auditory ERP attenuation during action observation. *Cognitive, Affective, & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-023-01083-7

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience*, *33*, 89-108. https://doi.org/10.1146/annurev-neuro-060909-153135

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for Sensory Prediction Deficits in Schizophrenia. *American Journal of Psychiatry*, *162*(12), 2384-2386. https://doi.org/10.1176/appi.ajp.162.12.2384

Shergill, S. S., White, T. P., Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2014). Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA Psychiatry*, *71*(1), 28-35. https://doi.org/10.1001/jamapsychiatry.2013.2974

Shulruf, B., Alesi, M., Ciochină, L., Faria, L., Hattie, J., Hong, F., Pepi, A.-M., & Watkins, D. (2011). Measuring collectivism and individualism in the third millennium. *Social Behavior and Personality: an international journal*, *39*(2), 173-187. https://doi.org/10.2224/sbp.2011.39.2.173

Shulruf, B., Hattie, J., & Dixon, R. (2007). Development of a New Measurement Tool for Individualism and Collectivism. *Journal of Psychoeducational Assessment*, *25*(4), 385-401. https://doi.org/10.1177/0734282906298992

Singelis, T. M. (1994). The Measurement of Independent and Interdependent Self-Construals. *Personality and Social Psychology Bulletin*, *20*(5), 580-591. https://doi.org/10.1177/0146167294205014

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. https://doi.org/10.1007/s00221-012-3204-7

Straube, B., van Kemenade, B. M., Arikan, B. E., Fiehler, K., Leube, D. T., Harris, L. R., & Kircher, T. (2017). Predicting the Multisensory Consequences of One's Own Action: BOLD Suppression in Auditory and Visual Cortices. *PLOS ONE*, *12*(1), e0169131. https://doi.org/10.1371/journal.pone.0169131

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219-239. https://doi.org/10.1016/j.concog.2007.03.010

Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: an interplay between prediction and postdiction. *Frontiers in Psychology*, *4*, 127. https://doi.org/10.3389/fpsyg.2013.00127

Tanaka, T., Matsumoto, T., Hayashi, S., Takagi, S., & Kawabata, H. (2019). What Makes Action and Outcome Temporally Close to Each Other: A Systematic Review and Meta-Analysis of Temporal Binding. *Timing & Time Perception*, *7*(3), 189-218. https://doi.org/10.1163/22134468-20191150

Tang, X., Zhu, W., Liu, X., Wang, T., Guan, J., & Dang, X. (2025). Motor Imagery Brain-Computer Interaction Using Different Modalities of Action Observation. *IEEE Transactions on Instrumentation and Measurement*, *74*, 1-11. https://doi.org/10.1109/tim.2025.3548181

Tast, V., Schroger, E., & Widmann, A. (2024). Suppression and omission effects in auditory predictive processing-Two of the same? *European Journal of Neuroscience*, *60*(2), 4049-4062. https://doi.org/10.1111/ejn.16393

Tast, V., Schroger, E., & Widmann, A. (2025). Auditory N1 Suppression and Omission N1 Do Not Share a Common Underlying Mechanism. *Psychophysiology*, *62*(6), e70094. https://doi.org/10.1111/psyp.70094

Teufel, C., Kingdon, A., Ingram, J. N., Wolpert, D. M., & Fletcher, P. C. (2010). Deficits in sensory prediction are related to delusional ideation in healthy individuals.

Neuropsychologia, *48*(14), 4169-4172. https://doi.org/10.1016/j.neuropsychologia.2010.10.024

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*(7), 1481-1489. https://doi.org/10.1162/jocn_a_00552

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), 1-11. https://doi.org/10.1186/1471-2202-14-2

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. https://doi.org/10.1016/j.cortex.2016.03.018

Turella, L., Erb, M., Grodd, W., & Castiello, U. (2009). Visual features of an observed agent do not modulate human brain activity during action observation. *NeuroImage*, *46*(3), 844-853. https://doi.org/10.1016/j.neuroimage.2009.03.002

Urgesi, C., Candidi, M., & Avenanti, A. (2014). Neuroanatomical substrates of action perception and understanding: an anatomic likelihood estimation meta-analysis of lesion-symptom mapping studies in brain injured patients. *Frontiers in Human Neuroscience*, *8*, 344. https://doi.org/10.3389/fnhum.2014.00344

van Elk, M., Salomon, R., Kannape, O., & Blanke, O. (2014). Suppression of the N1 auditory evoked potential for sounds generated by the upper and lower limbs. *Biological Psychology*, *102*, 108-117. https://doi.org/10.1016/j.biopsycho.2014.06.007

van Laarhoven, T., Stekelenburg, J. J., Eussen, M. L. J. M., & Vroomen, J. (2019). Electrophysiological Alterations in Motor-Auditory Predictive Coding in Autism Spectrum Disorder. *Autism Research*, *12*(4), 589-599. https://doi.org/10.1002/aur.2087

van Laarhoven, T., Stekelenburg, J. J., & Vroomen, J. (2021). Suppression of the auditory N1 by visual anticipatory motion is modulated by temporal and identity predictability. *Psychophysiology*, *58*(3). https://doi.org/10.1111/psyp.13749

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, *48*(3), 564-584. https://doi.org/10.1016/j.neuroimage.2009.06.009

Vroomen, J., & Stekelenburg, J. J. (2010). Visual Anticipatory Information Modulates Multisensory Interactions of Artificial Audiovisual Stimuli. *Journal of Cognitive Neuroscience*, *22*(7), 1583-1596. https://doi.org/10.1162/jocn.2009.21308

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., & Weil, L. (2006). Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). *Personality and Individual Differences*, *41*(5), 929-940. https://doi.org/10.1016/j.paid.2006.03.017

Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary Observations on Tickling Oneself. *Nature*, *230*(5296), 589-599. https://doi.org/10.1038/230598a0

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011a). The self in action effects: Selective attenuation of self-generated sounds. *Cognition*, *121*(2), 207-218. https://doi.org/10.1016/j.cognition.2011.06.011

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011b). The Self in Social Interactions: Sensory Attenuation of Auditory Action Effects Is Stronger in Interactions with Others. *PLOS ONE*, *6*(7), 1-3. https://doi.org/10.1371/journal.pone.0022723

Weiss, C., & Schutz-Bosbach, S. (2012). Vicarious action preparation does not result in sensory attenuation of auditory action effects. *Consciousness and Cognition*, *21*(4), 1654-1661. https://doi.org/10.1016/j.concog.2012.08.010

Weller, L., Schwarz, K. A., Kunde, W., & Pfister, R. (2017). Was it me? – Filling the interval between action and effects increases agency but not sensory attenuation. *Biological Psychology*, *123*, 241-249. https://doi.org/10.1016/j.biopsycho.2016.12.015

Wen, W. (2019). Does delay in feedback diminish sense of agency? A review. *Consciousness and Cognition*, *73*, 102759. https://doi.org/10.1016/j.concog.2019.05.007

Whitford, T. J., Mathalon, D. H., Shenton, M. E., Roach, B. J., Bammer, R., Adcock, R. A., Bouix, S., Kubicki, M., De Siebenthal, J., Rausch, A. C., Schneiderman, J. S., & Ford, J. M. (2011). Electrophysiological and diffusion tensor imaging evidence of delayed corollary discharges in patients with schizophrenia. *Psychological Medicine*, *41*(5), 959-969. https://doi.org/10.1017/S0033291710001376

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 593-602. https://doi.org/10.1098/rstb.2002.1238

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729-R732. https://doi.org/10.1016/S0960-9822(01)00432-8

Woods, D. L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalography and Clinical Neurophysiology. Supplement*, *44*, 102-109.

Yamada, A.-M., & Singelis, T. M. (1999). Biculturalism and self-construal. *International Journal of Intercultural Relations*, *23*(5), 697-709. https://doi.org/10.1016/S0147-1767(99)00016-4

5 Eidesstattliche Erklärung

Eidesstattliche Erklärung gemäß §5 der Promotionsordnung vom 15.6.2018 der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf:

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist. Die Dissertation wurde in der vorliegenden oder ähnlichen Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, 05.08.2025
Datum

Sophie Egan
Sophie Egan

6 Appendix A: Original Research Articles

In this section, the original articles of the three studies are presented. Specifically, the published article of Study 1, the submitted manuscript of Study 2, and the published article of Study 3 are included on the following pages after the contribution statements.

Study 1 has been published in Biological Psychology with the following reference:

Egan, S., Ghio, M., & Bellebaum, C. (2023). Auditory N1 and P2 attenuation in action observation: An event-related potential study considering effects of temporal predictability and individualism. *Biological Psychology*, *180*, 108575. https://doi.org/10.1016/j.biopsycho.2023.108575

I wrote the original draft and integrated revisions from the co-authors. I contributed to conceptualisation and methodology. I conducted and supervised the data acquisition. I wrote part of the code for data processing, conducted the formal analysis with support, interpreted the results, created visualisations and worked on the data curation.

The second study has been submitted for publication:

Egan, S., Weber, C., Ghio, M., & Bellebaum, C. (2025). I, You, Robot: Attenuation for Auditory Outcomes of Actions Performed by Different Agents Shows Distinct Patterns for N1 and P2 Amplitudes [Manuscript submitted for publication].

I wrote the original draft and integrated revisions from the co-authors. I contributed to the conceptualisation and the methodology, including the creation of experimental stimuli. I supervised the data acquisition. I wrote the code for data processing, conducted the formal analysis, interpreted the results, created visualisations, and worked on the data curation.

The third study has been published in the Journal of Cognitive Neuroscience with the following reference:

Egan, S., Seidel, A., Weber, C., Ghio, M., & Bellebaum, C. (2024). Fifty Percent of the Time, Tones Come Every Time: Stronger Prediction Error Effects on Neurophysiological Sensory Attenuation for Self-generated Tones. *Journal of Cognitive Neuroscience*, *36*(10), 2067-2083. https://doi.org/10.1162/jocn_a_02226

I wrote the original draft and integrated revisions from the co-authors. I supervised the data acquisition. I adapted code for data processing, conducted the formal analysis with support, interpreted the results, created visualisations, and worked on the data curation.

6.1 Study 1

Biological Psychology 180 (2023) 108575

Contents lists available at ScienceDirect

Biological Psychology

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

Auditory N1 and P2 attenuation in action observation: An event-related potential study considering effects of temporal predictability and individualism

Sophie Egan [*], Marta Ghio, Christian Bellebaum

Institute of Experimental Psychology, Heinrich Heine University, Düsseldorf, Germany

ARTICLE INFO

Keywords:
Auditory sensory attenuation
Action observation
EEG
Temporal predictability
Individualism-collectivism

ABSTRACT

Tones that are generated by self-performed actions elicit attenuated N1 and P2 amplitudes, as measured by electroencephalography (EEG), compared to identical external tones, which is referred to as neurophysiological sensory attenuation (SA). At the same time, self-generated tones are perceived as less loud compared to external tones (perceptual SA). Action observation led in part to a similar neurophysiological and perceptual SA. The perceptual SA in observers was found in comparison to tones that were temporally predictable, and one study suggested that perceptual SA in observers might depend on the cultural dimension of individualism. In this study, we examined neurophysiological SA for tones elicited by self-performed and observed actions during simultaneous EEG acquisitions in two participants, extending the paradigm with a visual cue condition controlling for effects of temporal predictability. Moreover, we investigated the effect of individualism on neurophysiological SA in action observation. Relative to un-cued external tones, the N1 was only descriptively reduced for tones that were elicited by self-performed or observed actions and significantly attenuated for cued external tones. A P2 attenuation effect relative to un-cued external tones was found in all three conditions, with stronger effects for self- and other-generated tones than for cued external tones. We found no evidence for an effect of individualism. These findings add to previous evidence for neurophysiological SA in action performance and observation with a paradigm well-controlled for the effect of predictability and individualism, showing differential effects of the former on the N1 and P2 components, and no effect of the latter.

1. Introduction

Sensory attenuation (SA) describes an effect that a voluntary action has on the processing of its sensory consequences (Jack et al., 2021; Timm et al., 2014). In an example of auditory processing, when I generate a tone by pressing a button, I perceive it as quieter and the neurophysiological response to the tone in my brain is reduced compared to an external tone (i.e., a tone I did not generate; Schafer & Marcus, 1973; Weiss et al., 2011). These two phenomena are possibly linked and were described separately in the literature. In the following, we use the terms "perceptual SA" and "neurophysiological SA" to distinguish them, and "SA" to refer to both. Acknowledged as the result of predictive mechanisms (but see Horváth et al., 2012), there are different approaches to explaining SA. On the one hand, internal forward models based on a copy of the motor command (Wolpert & Flanagan, 2001) were used to explain the SA effect. In this context, the cerebellum was proposed to be involved in generating motor-based sensory predictions (Blakemore et al., 2001; Blakemore & Sirigu, 2003; Knolle et al., 2013). On the other hand, it was suggested that the predictive mechanisms underlying the attenuation of self-generated tones might be more general, i.e., motor-independent and not necessarily intention-based, and apply also to external tones that are similarly predictable (Dogge et al., 2019; Kiepe et al., 2021; Korka et al., 2022). Interestingly, deficits in predictive mechanisms underlying SA and in agency attribution have, for example, been found in patients with psychotic disorders, such as schizophrenia (Ford et al., 2007; Ford et al., 2001; Krugwasser et al., 2022; Whitford, 2019).

In humans, neurophysiological SA for auditory stimuli has frequently been investigated with electroencephalography (EEG), which is well suited to explore the temporal dynamics of auditory processing. More specifically, the amplitudes of the N1 and P2 event-related potential (ERP) components have been found to be reduced for self-generated

---

[*] Correspondence to: Universitätsstraße 1, 40225 Düsseldorf, Germany.
E-mail address: Sophie.Egan@hhu.de (S. Egan).

https://doi.org/10.1016/j.biopsycho.2023.108575
Received 16 February 2023; Received in revised form 30 April 2023; Accepted 4 May 2023

relative to external tones (e.g., McCarthy & Donchin, 1976; Schafer & Marcus, 1973). However, there are findings suggesting that these components are differentially modulated in SA, thus suggesting that they might represent different processes (Korka et al., 2022). N1 attenuation was reduced in patients with cerebellar lesions compared to controls, while the P2 was not affected (Knolle et al., 2013), suggesting that only the N1 might reflect cerebellar prediction mechanisms. This was in part supported by Cao et al. (2017), who found the N1 and P2 counterparts in magnetencephalography (M100 and M200) to be differently influenced by cerebellar transcranial magnetic stimulation. Further, the P2 but not the N1 was correlated with the experience of agency (Timm et al., 2016) and was affected by the induction of high or low control illusions over the elicitation of an identical tone in a preceding task (Seidel et al., 2021). Therefore, the P2 was suggested to reflect a retrospective, more conscious agency judgement (Knolle et al., 2013; Timm et al., 2016). In line with this, increasing the action-outcome delay in a small percentage of trials selectively affected P2 amplitudes, in that an increase of delay reduced attenuation (Pinheiro et al., 2019), possibly due to an increase of uncertainty about the agency.

1.1. Temporal predictability

In the standard self-generation paradigm applied to investigate neurophysiological SA, the comparison condition to self-generated tones consists of a sequence of external un-cued tones (for a review see Horváth, 2015), and it has been argued that the difference between the ERP amplitudes in the two conditions might therefore be explained simply by the higher temporal predictability of the first (Dogge et al., 2019; Kiepe et al., 2021). Temporal predictability refers to the ability to predict the onset of a stimulus (Hughes et al., 2013). If general prediction mechanisms lead to the neurophysiological SA effect, the amplitudes elicited by tones that are temporally predictable based on non-motor information (e.g., a visual cue) should be attenuated relative to un-cued tones as well.

Previous studies provided controversial evidence on the effects of temporal predictability: There are studies that showed neurophysiological SA for tones that were temporally cued (i.e., tones came in a highly regular interval) or cued (i.e., preceded) by visual stimuli (Kaiser & Schütz-Bosbach, 2018; Schafer & Marcus, 1973; Sowman et al., 2012). This was found for the N1 (Kaiser & Schütz-Bosbach, 2018) and for both the N1 and P2 component (Schafer & Marcus, 1973; Sowman et al., 2012). Interestingly, however, Schafer and Marcus (1973) already found that the N1, but not the P2, attenuation for self-generated stimuli was stronger than for temporally cued ones. This is in line with recent studies reporting that N1 amplitudes of self-generated tones were still attenuated in comparison to visually cued tones (Klaffehn et al., 2019; Lange, 2011). In the study by Lange (2011), N1 amplitudes of self-generated tones and tones preceded by a cue with either constant or variable delays were compared, and a reduction was found for self-generated tones relative to cued tones irrespective of the length (350 ms vs. 750 ms) or type (constant or variable) of delay. In the study by Klaffehn et al. (2019), ERPs of self-generated tones with and without a delay of 750 ms between button press and tone were compared to externally generated visually cued (750 ms between cue onset and tone) and un-cued tones. N1 amplitudes were attenuated for self-generated tones relative to cued and un-cued tones with and without delay, and P2 amplitudes were attenuated for self-generated tones only in the condition without the delay.

Taken together, these findings suggest that a stronger attenuation occurs for self-generated tones than temporally predictable external tones (with a potential role of the delay between the button press or preceding cue on the one hand and the tone on the other hand), which could reflect the higher reliability of the own action for temporal predictability and/or a specific action-related process. However, SA was also found for cued relative to un-cued tones, suggesting that SA for self-generated tones may be the result of a combination of forward model or intention-based and more general predictions (Korka et al., 2022).

1.2. Sensory attenuation and action observation

Studying SA in action observation is particularly interesting, as imaging studies show an overlap of neural activity when performing an action and when observing the same action being performed by someone else (e.g., Buccino et al., 2013). Hence, predictions regarding the consequences of observed actions may be similar to predictions based on self-generated actions and thus be action-specific (Wolpert et al., 2003). Alternatively, predictions based on observed actions may be similar to predictions based on visual cues (e.g., Kaiser & Schütz-Bosbach, 2018; Lange, 2011) and thus only be dependent on temporal predictability (Korka et al., 2022).

A few studies have investigated a potential perceptual SA effect on the perceived loudness of tones in action observation (Cao & Gross, 2015; Sato, 2008; Weiss et al., 2011). In these studies, a second tone was played in variable loudness after an action-generated or cued tone and participants made a loudness judgment. For the observation condition, in which the experimenter generated the tones, the results were not consistent: Sato (2008) found perceptual SA in observers, whereas Weiss et al. (2011) did not. Cao and Gross (2015) noted that the participants of the first study were Japanese and the ones of the second were German. They sought to explain the discrepancy in findings on perceptual SA in observers as an effect of culture and indeed found culture dependent differences: In their Chinese sample there was perceptual SA for action observation but not in their British sample (Cao & Gross, 2015).

Cultures can differ in several dimensions, the one of interest in the behavioural study being the individualism-collectivism dimension (Cao & Gross, 2015; Hofstede, 2011). This dimension describes a range of different aspects but can be summarised as on the one end, individualistic people valuing their personal independence and on the other end, collectivistic people valuing belonging to their group (Hofstede, 2011; Shulruf et al., 2007). Individualism and collectivism can be seen as two distinct aspects of personality at the individual level (Hofstede, 2011). Indeed, questionnaires usually feature two scores and participants were found to be individualistic or collectivistic (i.e., scored high in one and low in the other), bicultural (i.e., scored high in both), or a-cultural (i.e., scored low in both; Yamada & Singelis, 1999). The results of Cao and Gross (2015) suggested that individualism, and not collectivism, modulated the cultural difference in the perceptual SA for observers: The less individualistic the participant was, the greater the perceptual SA in observation.

To investigate the neurophysiological SA in observers, Ghio et al. (2017) introduced an observation condition to the self-generation paradigm, in which participants viewed a picture sequence showing another person's button press from the first-person perspective. Neurophysiological SA was found in observers in comparison to external tones (presented in a separate experimental block as in the standard self-generation paradigm) for the N1 and the P2, though the latter was weaker than the SA effect for tones generated by own actions (Ghio et al., 2017). In a following study, the observing participant sat next to the performing participant and EEG was acquired in both participants simultaneously (Ghio et al., 2021). In contrast to the previous study, N1 amplitudes of other-generated tones during action observation were not attenuated, while P2 amplitudes of other-generated tones were attenuated and were not significantly different to P2 amplitudes of self-generated tones (Ghio et al., 2021).

The differences in N1 findings between the studies might be attributed to the different delays between observed button press and tone: The longer delay (about 125 ms) in the first study (Ghio et al., 2017) compared to the one in the second study (about 50 ms; Ghio et al., 2021) may have allowed the visual input to be processed in time and thus enabled temporal predictability of the tone and neurophysiological SA for N1 amplitudes in observers (see Ghio et al., 2021, for a more detailed discussion), as it has been shown for visually cued tones (e.g, Kaiser &

Schütz-Bosbach, 2018). Both studies lacked a condition controlling for temporal predictability (Ghio et al., 2021; Ghio et al., 2017). This raises the question whether the neurophysiological SA effects for action observation go beyond those reported for tones that are cued by another visual cue (e.g., Kaiser & Schütz-Bosbach, 2018).

Moreover, the potential effect of individualism may have contributed to the inconsistency in the SA findings. It should be noted that both observation studies on the neurophysiological SA acquired a German sample (Ghio et al., 2021; Ghio et al., 2017), like the study of Weiss et al. (2011). Therefore, while individualism possibly influenced the perceptual SA in observation (Cao & Gross, 2015; Sato, 2008; Weiss et al., 2011), a potential effect of individualism on the attenuation of ERP components (P2 and partly N1) of other-generated tones has not been investigated yet.

1.3. *Present study*

In the present study two research questions were addressed. First, we wanted to examine whether a neurophysiological SA effect, as reflected in N1 and P2 amplitudes, that goes beyond an unspecific effect of predictability can be found for action observation. We thus used the observation version of the self-generation paradigm (Ghio et al., 2021; Ghio et al., 2017) and added a cue condition to account for the effect of temporal predictability (Hughes et al., 2013; Klaffehn et al., 2019). To study the neurophysiological SA effect for observed action outcomes we applied a simultaneous setting for the acquisition (see Ghio et al., 2021). In the present study the two participants switched the roles of performer and observer during the same EEG acquisition session, and both saw the actions from a first-person perspective (Ghio et al., 2017). All participants were thus presented with four experimental conditions, in which the tone generation and cueing was manipulated. Being physically identical in all conditions, the tone was either self-generated (self), generated by an observed person (other), preceded by a visual stimulus (cued) or not preceded by an event (un-cued). Both the observed action and the cue can be considered visual stimuli that precede the tone, with the difference that in the observation condition it is a human action and this action is the same as for self-generation. Importantly, a 350 ms delay between the preceding event and the tone was applied throughout, enabling predictions to be formed based on the visual input in the other and cued conditions (e.g., Kaiser & Schütz-Bosbach, 2018; Lange, 2011).

The second research question was related to a potential effect of individualism on the neurophysiological SA effect in action observation, based on the behavioural study by Cao and Gross (2015). For this purpose, we applied the Auckland Individualism Collectivism Scale (AICS; Shulruf et al., 2011), which includes several subscales: Individualism consisted of responsibility, uniqueness, and competitiveness; collectivism consisted of advice and harmony (Shulruf et al., 2007). We aimed to acquire a sample with varying degrees of individualism and to use the individualism score as a continuous variable in the analysis. Cao and Gross (2015) further considered delusional ideation and empathy. They found that delusional ideation was related to reduced SA in action performance (only for the British sample) but empathy did not relate to SA in action observation. Similarly, we added one questionnaire to account for a potential influence of anomalous perception (Bell et al., 2006), which is more directly relevant to tone perception than items measuring delusional ideation (e.g., beliefs), and one questionnaire to account for potential effects of empathy (Lawrence et al., 2004) on SA in action performance and observation.

Concerning the first research question we expected a significant effect of tone trial type. For self-generated tones, we expected an attenuation of the amplitudes of both N1 and P2 compared to external un-cued tones, as it has been shown in several previous studies (Horváth, 2015), as well as to cued tones (Klaffehn et al., 2019). For action observation, given the perspective and the action-outcome delay in our study, the N1 of other-generated tones should be attenuated compared to un-cued tones (see Ghio et al., 2017; Seidel et al., 2023). Compared to cued tones, there are two potential scenarios: If the neurophysiological SA in observation is action-specific (Wolpert et al., 2003), a stronger N1 attenuation should occur for other-generated vs. cued tones. However, if prediction based on observation is comparable to prediction based on other visual cue stimuli (and thus depends only on temporal predictability), there should not be a significant difference between the processing of other-generated and cued tones. The interpretation of the P2 attenuation in terms of agency attribution (e.g., Timm et al., 2016) would suggest action-specificity (Ghio et al., 2021; Ghio et al., 2017). Therefore, a reduction of the observer P2 amplitude might be found not only compared to un-cued but also compared to cued tones.

Concerning the second research question, the neurophysiological SA for action observation may only be found in participants with low individualism, as suggested by Cao and Gross (2015) for the perceptual SA in observers. Then, the interaction of tone trial type and individualism would be significant, in that only amplitudes for other-generated tones are correlated with individualism scores. It is unclear whether the N1 or P2 attenuation or both might correspond to the perceptual SA and the interaction with individualism could be significant only for one component. As the subscales of the AICS, instead of the higher-order scores, differentiated the tested ethnic groups, this could further pertain to only a subset of the three individualism subscales (Shulruf et al., 2007).

2. Method

2.1. *Sample*

Based on our previous work on neurophysiological SA in action and observation by applying simultaneous EEG acquisitions in pairs of participants (Ghio et al., 2021) and on our work on relating continuous measures reflecting participants' trait characteristics to neurophysiological data (Bellebaum et al., 2020), we aimed to recruit 60 participants, also considering a potential drop-out rate of 10–15%. The acquisition was planned in pairs of participants (see below for details) and we thus aimed to test 30 pairs. Some of the invited participants did not show up, however, so that an experimenter had to jump in as the second person in a pair. We thus stopped recruiting after 57 participants in total had been tested. Participants with a history of neurological or psychiatric disorders or medication use were not considered for this study. Moreover, eight participants were excluded from data analysis: two due to technical problems and six due to too few trials for the analysis of the EEG data (see below). The final sample used in the analysis thus consisted of 49 participants (39 female). The mean age was 24.65 years ($SD = 5.30$). We recruited some participants with different nationalities to increase the chance of variance in the individualism scores relevant to address our second research question. Most participants were German (39), three of which stated to have a second nationality (Kazakhstan, Poland, USA). The rest of the participants were Bulgarian (one), French (one), Hungarian (one), Romanian (one), Russian (one), and Vietnamese (five). Each participant gave written informed consent and received up to 35 euros or course credit for their participation. The ethics committee of the Faculty of Mathematics and Natural Sciences at Heinrich Heine University has approved this study, which was conducted in accordance with the ethical standard of the declaration of Helsinki.

2.2. *Experimental procedure*

2.2.1. *Questionnaires*

Participant information and a series of questionnaires were given online via Soscisurvey before participation in the EEG experiment (self-generation paradigm). First was a demographic questionnaire including questions on nationality. Further, three questionnaires were presented in random order. One questionnaire was the AICS (Shulruf et al., 2011) as measurement of individualism (and collectivism). To account for potential effects of anomalous perception and empathy on the neurophysiological SA (in observation), the Cardiff Anomalous Perception Scale (CAPS; Bell et al., 2006) and the Empathy Quotient (EQ; Lawrence et al., 2004) were added as well (Cao & Gross, 2015). We regularly checked the distribution of AICS individualism scores in our participants during data acquisition to see that they varied on this scale.

2.2.2. *EEG experiment*

The EEG acquisition was done in pairs (see Fig. 1A), and 25 pairs consisted of two participants each, while in seven cases, only one participant came to the acquisition and an experimenter took part as second person of the pair (this was done by four different experimenters: two took over once, one twice, and one three times). The EEG experiment started with a learning phase, followed by an experimental phase.

2.2.3. *Learning phase*

After giving written informed consent, participants began the learning phase of the self-generation paradigm that consisted of two parts. Its purpose was to train participants to press a button in a regular fixed interval and to enable participants to learn the association between (own and observed) button press and tone. In the first part, participants were presented a tone every 2.4 s and they were instructed to press a button whenever the tone appeared. For each of the two participants this was done in two blocks consisting of 30 trials each, one block using the right hand and one block using the left hand, respectively. Thus, each participant actively pressed the button in two blocks and observed the other participant pressing the button in the other two. Block order was fixed due to the constraints of the setup (see Fig. 1A).

In the second part of the learning phase, participants pressed the button at will, thereby eliciting a tone with a 350 ms delay. They were asked to press the button in the interval, in which the tone was presented in the first part of the training (2.4 s) and received visual feedback for when the intervals were too short (< 1.8 s) or too long (> 3 s). Again, there were four blocks, with each participant being performer in two blocks (actions performed with left and right hand) and observer in the other two, and the order was fixed. Blocks in the second part of the learning phase consisted of 60 trials each and began with the presentation of three tones in the 2.4 s interval to remind the participants of the learned interval for button presses. After the learning phase, the participants were prepared for the EEG acquisition.

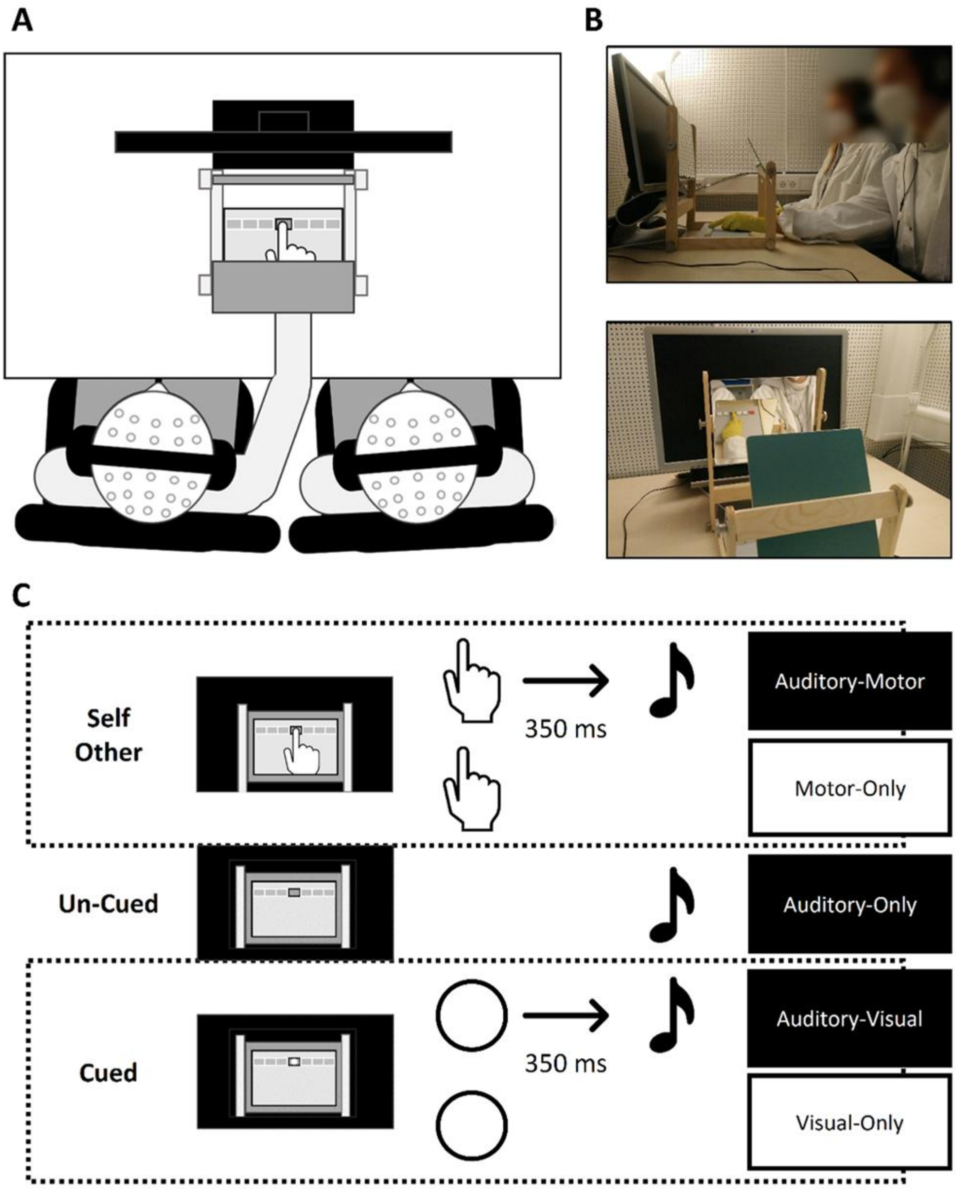


Fig. 1. *Setup and Experimental Conditions.* (A) The experimental setup in a block that required button presses (auditory-motor or motor-only). For the left participant the setup would be for the self condition, for the right participant for the other condition. The mirror construction was placed above the response box in these blocks. The participant on the left could comfortably press the button only with the right hand, and the participant on the right could press the button with the left hand. Therefore, seating position had to be switched once the right-handed block for one and left-handed block for the other were completed, before the same condition was run with the respective other hand. (B) Photos depicting the view from the side (top) and from the observing participant's point of view (bottom). Participants wore masks (according to the pandemic guidelines) and gloves of different colour to clearly differentiate between participants' hands. (C) Experimental conditions. Enclosed in the dotted lines are the conditions for which ERPs were subtracted later in the analysis. The condition used for correction is in the white box, respectively. Visually, the participants had the same input: In auditory-motor and motor-only blocks they saw the hand reflected in the mirror, in the auditory-only block they saw a picture of the response box (as reflected in the mirror) on the screen and in auditory-visual and visual-only blocks they saw the same picture with a white circle as visual cue on the response button in the picture.

During both parts of the learning phase and irrespective of whether they actively performed or observed the button press actions, participants were asked to look at the button press action in a mirror that projected the image of the performer's hand and the button box in front of the participants at the height of a computer screen (see Fig. 1B). With this change in the experimental procedure compared to our previous study applying a simultaneous EEG acquisition with a performer and an observer (Ghio et al., 2021) we aimed to ensure that visual stimulation and posture of both participants were comparable.

2.2.4. Experimental phase

For each of the self, other, and cued conditions we included (a) one block, in which tones followed the preceding event (i.e., the [observed] action or the visual cue) in all trials, and (b) one block in which all trials were without tones, which served to control for ERPs related to the event preceding the tone (see Fig. 1C). Specifically, for the self and other conditions we included (a) the auditory-motor block, during which one participant of the pair pressed the button as learned during the learning phase and thereby generated a tone that followed with a delay of 350 ms after button press onset in each trial; (b) the motor-only block, during which in each trial one of the participants pressed the button as learned but no tone was generated. In both blocks, the other participant was asked to observe the button presses. Both participants looked at the button press actions in the mirror, as they did during the learning phase (see Fig. 1 for the setup in these blocks).

For the cued condition we included (a) the auditory-visual block, during which the tone trial sequence as generated in the preceding auditory-motor block was presented. Visually a picture, which showed the response box reflected in the mirror, was presented on the computer screen and participants were asked to fixate the button that they pressed in the blocks involving actions. In each trial, a tone was preceded by a visual cue (i.e., a white circle) displayed on the button 350 ms before the tone onset. We included (b) the visual-only block, during which the visual cues were presented without the tones. In both blocks, each cue presentation was matched to the respective button press in the preceding auditory-motor (for the auditory-visual) or motor-only (for the visual-only) block with respect to onset and duration.

For the un-cued condition, we included the auditory-only block, during which the tone trial sequence as generated in the preceding auditory-motor block was presented without any preceding event. As in the visual blocks, participants were asked to fixate the button of the response box in the picture on the screen, but no visual cue appeared before the tone (see Fig. 1C for the sequence of events in the different conditions).

Each block consisted of two subblocks, with a total of 120 trials. In the blocks that required button presses from one of the participants, 60 trials were performed with the right hand and 60 trials with the left hand. Each subblock, including those that did not feature tones, began with three tones in the 2.4 s interval to serve as reminder of the learned interval. Due to the constraints of the setup and the dependency between blocks, we created four versions of the experiment, each with a different but fixed subblock order.

2.3. Experimental setup

The two participants were seated next to each other with the response box in between (Response Pad Model RB-740 from Cedrus Corporation, www.cedrus.com), in front of a 22-inch LCD computer screen with a screen resolution of 1680 * 1050 and a 60 Hz refresh rate (see Fig. 1A). For the auditory-motor and motor-only blocks, a mirror construction was placed on top of the response box by an experimenter. The construction had two mirrors placed so that the response box was not visible for the participants and the reflection of it could be seen at the height of the computer screen (Fig. 1B). The software Presentation (version 20.3, Build 02.25.19) from Neurobehavioral Systems, Inc. (http://www.neruobs.com) was used for the presentation of all stimuli. The tone was a 680 Hz tone that lasted 50 ms, with an increase of amplitude in the first and a decrease in the last 5 ms, and was delivered via Sennheiser (HD 201) headphones.

2.4. EEG recording

We followed the same procedure as in Ghio et al. (2021) for the simultaneous EEG acquisition. First, we attached two reference electrodes on the mastoids (TP9, TP10) of the participants. Then, each participant was fitted with an Easycap electrode cap (www.easycap.de), which held 28 passive Ag/AgCl electrodes at F7, F3, Fz, F4, F8, FT7, FC4, FCz, FC4, FC8, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, PO7, PO3, POz, PO4, PO8, and AFz for the ground electrode. In the case of five participants, however, data from up to five parietal and parieto-occipital electrodes were not acquired due to technical problems during data acquisition. An electrode under the right eye and electrodes at F9, F10, and FP2 were used to record eye movements. BrainVision Recorder (Version 1.21.0402, Brain Products GmbH, www.brain-products.com) and BrainAmp DC-32 channel AC/DC amplifiers (one for each participant; one of them MR plus) were used to record the EEG data (1000 Hz sampling rate, 0 Hz low cut-off, 1000 Hz high cut-off, Notch- and other online software filters disabled) and to apply direct-current corrections online. Impedances were kept below 10 kΩ.

2.5. Data analysis

2.5.1. Questionnaires

The AICS individualism and collectivism scores were calculated as the means of individualism and collectivism items (1–6 scale), respectively (Shulruf et al., 2011). The sub-scores, likewise, were calculated as the means of the respective items. The sub-scores of the individualism scale were competitiveness (seven items), uniqueness (four items), and responsibility (four items). The sub-scores of the collectivism scale were advice (seven items) and harmony (four items).

The CAPS total score was calculated as the sum of yes-answers to the 32 items (Bell et al., 2006). For all items with yes-answers the participants further rated the distress, intrusiveness, and frequency of the sensation (1–5 scale). The sums of the ratings made up these three CAPS scores. All CAPS scores were highly skewed (>1.10). For the EQ score, we calculated the sum of the 60 items (4-point scale valued 0, 0, 1, and 2 after ratings were flipped for reverse items; Lawrence et al., 2004).

2.5.2. EEG pre-processing

The data was pre-processed with BrainVision Analyser (Version 2.20.7383, Brain Products): After a direct current detrend, Butterworth zero phase filters (low cut-off = 0.1 Hz, 12 dB/oct; high cut-off = 30 Hz, 12 dB/oct), Notch-Filter (50 Hz), and ocular correction independent component analysis (steps = 512, Infomax restricted biased) to remove blinking and eye movement artifacts, the data was segmented based on tone markers in the auditory-motor, auditory-visual and auditory-only blocks. The segments began 150 ms before the tone and were 600 ms long in total. The corresponding time window was segmented for the blocks not containing tones (motor-only and visual-only), based on the markers of the button press or the visual cue, respectively. Due to technical reasons the delays between events (i.e., button press or visual cue) and tones did not equal 350 ms exactly in all participants. Moreover, for cued tones the event-tone delay was systematically shorter, as the presentation of the cue was delayed till the next frame could be shown on the computer monitor. To account for this, the median event-tone delay for each participant was calculated and these medians were used to calculate the corresponding time window in the motor-only and visual-only blocks. In all blocks, trials with responses shorter than 1.8 s and longer than 10 s, as well as double presses were excluded. The 100 ms before the tone (and the corresponding time window in the correction blocks) served as baseline. Before averaging across trials, we discarded segments with artifacts as identified with the automatic procedure in BrainVision Analyser based on the following parameters: maximal allowed voltage step of 50 µV/ms, maximal allowed difference of values of 100 µV in 100 ms intervals, allowed amplitudes between − 100 and 100 µV, and lowest allowed activity of 0.5 µV in 100 ms. Six participants that had less than 60% of the segments in any block were excluded from analysis.

2.5.3. N1 and P2 amplitudes

Further processing and visualisation of the data was done in R (Version 4.2.2, 2022–10–31; R Core Team, 2022) with the ggplot2 package (Wickham, 2016). The average signal time-locked to the (observed) button presses in the motor-only blocks was subtracted from the averages for tones of the auditory-motor blocks to yield the corrected signal for the self and other conditions, respectively (see Ghio et al., 2021). Similarly, the average signal time-locked to the visual cue in the visual-only block was subtracted from the condition averages of the cued tones of the auditory-visual block, yielding the corrected cued tone condition. The labels "self, other, and cued conditions" refer to these corrected ERPs (for visualisation of the correction see Figure A1 in the supplement). A paired sample *t*-test revealed that there was no significant difference between the length of the intervals pressed in self-generation auditory-motor ($M = 2626$ ms, $SD = 342$) and motor-only blocks ($M = 2564$ ms, $SD = 363$), $t(48) = 1.792$, $p = .079$, $d = .256$. Although based on our experimental design the length of the intervals between button presses in the self- and other-generation condition should have in principle been the same, due to exclusions and participant-experimenter acquisitions the two datasets were partially different. Therefore, we checked the length of the intervals between button presses also in the other-generation condition. We found that the intervals were significantly longer in the other-generation auditory-motor blocks ($M = 2630$ ms, $SD = 380$) than in the other-generation motor-only blocks ($M = 2532$ ms, $SD = 345$), $t(48) = 2.693$, $p = .010$, $d = .385$. Despite this difference indicating that the correction might not be fully appropriate as other action-related processes could be involved (Horváth, 2015), the mean difference was relatively small (98 ms) and we assume that the correction can yet be applied.

Similar to previous studies, as we identified the electrode with the highest N1 and P2 peaks to be the FCz (see Fig. 2A; see Figure A1 in the supplement for a comparison of the ERPs of midline electrodes), the analysis was restricted to this electrode (e.g., Ghio et al., 2021; Neszmélyi & Horváth, 2017). We defined the N1 and P2 amplitudes as mean signal amplitudes in a 30 ms and 50 ms interval, respectively, around the respective peak in the grand average across all conditions according to the collapsed localiser technique (Luck & Gaspelin, 2017). The different interval lengths (i.e., 30 and 50 ms) for the means of N1 and P2 were chosen based on visual inspection of the peaks of the condition-specific averages (see Fig. 2B), which suggested different peak widths for the N1 and P2. The peaks in the grand average were defined as local minimum in the time interval between 50 and 150 ms after the tone for the N1, which yielded a peak at 107 ms (interval for means: 92–122 ms), and as local maximum in the time interval between 150 and 250 ms after the tone for the P2, which yielded a peak at 186 ms (interval for means: 161–211 ms).

2.5.4. Statistical analyses

Data were analysed by fitting linear mixed models and all analyses were done in R (Version 4.2.2, 2022–10–31; R Core Team, 2022). One advantage of this statistical analysis approach is the ease of analysing an interaction of a categorical and a continuous variable, which is relevant for our second experimental question.

To address our first experimental question, namely whether there is an SA effect for action observation that goes beyond an unspecific prediction effect, we specified models (one for the N1 and one for the P2) with the tone trial type as fixed effect (coded as 1–4 for un-cued, cued,

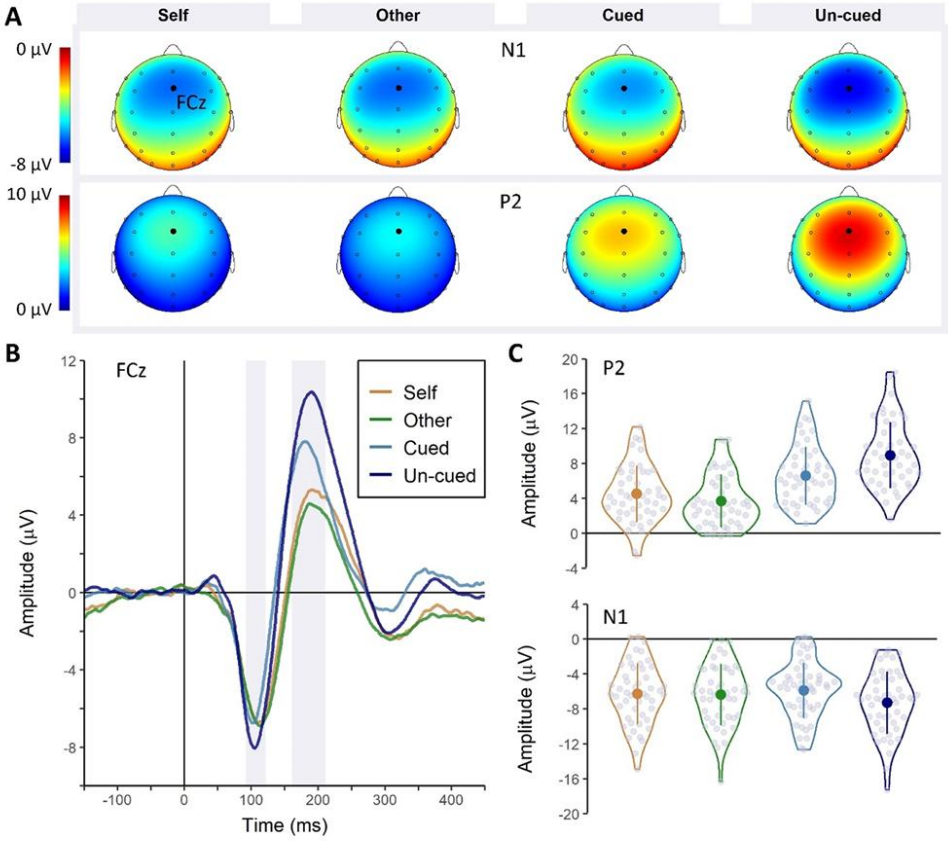


Fig. 2. *Topographies, ERPs and Amplitudes of N1 and P2.* (A) Topographies for the time interval of the N1 (top; 92–122 ms) and P2 for each condition (bottom; 161–211 ms) are depicted. (B) The averaged event-related potentials (ERPs) at FCz are displayed. See Fig. A1 in the supplement for individual ERPs. Please note that for the self, other and cued conditions the ERPs reflect difference waves, as the ERPs from the respective control conditions were already subtracted. (C) The mean amplitudes of the respective intervals for N1 and P2 in each condition are presented. Error bars represent one standard deviation. For a visualisation of the within-subject effect of tone trial type see Fig. A2 in the supplement, in which individual N1 and P2 amplitudes are connected for each pairwise comparison of the four conditions.

other, self). We modelled a random intercept by participant but not a random slope because for each ERP component (N1, P2) we entered into the analyses one mean amplitude value for each participant for each of the four conditions. This model we called the basic model: amplitude ~ tone trial type + (1|participant). The model was calculated using restricted maximum likelihood (Meteyard & Davies, 2020) as implemented in the lme4 package (Bates et al., 2015). Outliers were defined as having a Cook's distance over 4/n (with $n = 49$; Arimie et al., 2020; Cook, 2000) with the help of influence.ME (Nieuwenhuis et al., 2012). For the N1 amplitudes, no outlier was found and the sample consisted of the 49 participants described in the sample section. For the P2, two participants were excluded as outliers and thus the sample consisted of 47 participants (38 female; age: $M = 24.53$ years, $SD = 5.37$). For the decomposition of fixed-effects contributions we applied the anova function as implemented in the lme4 package (Bates et al., 2015). To obtain p-values in type III ANOVA and summary tables for the model fit we used lmerTest (Kuznetsova et al., 2017). The effect size was calculated with the package effectsize (Ben-Shachar et al., 2020). As suggested by Volpert-Esmond et al. (2021), we used the Satterthwaite approximation for the degrees of freedom. Bonferroni-corrected post-hoc pairwise comparisons were calculated with the package emmeans (Lenth, 2022).

To address our second experimental question and test for the influence of culture, particularly of individualism, first we standardised AICS scores (as suggested for continuous variables by Volpert-Esmond et al., 2021). We specified separate models starting from the basic model and adding one of the continuous variables derived from the applied questionnaire, which was modelled as main effect and in interaction with the tone trial type factor: amplitude ~ tone trial type * variable + (1| participant). As before, we entered one mean amplitude value for each participant for each of the four conditions into the analyses, separately for the ERP components (N1 and P2). The culture models were: One model including the individualism score, which was of main interest for our research question, one with the collectivism score, and one for each individualism sub-score (i.e., competitiveness, responsibility and uniqueness) and collectivism sub-score (i.e., advice, harmony). Using the package performance (Lüdecke et al., 2021) we compared the models with the basic model via likelihood ratio tests, as suggested for the comparison of nested models (Meteyard & Davies, 2020). For the model comparison, these models were calculated using maximum likelihood (Meteyard & Davies, 2020). To account for the potential effects of anomalous perception and empathy, we log-transformed and standardised CAPS scores and standardised EQ scores before specifying separate models including the main effect of each variable (i.e., CAPS total score and ratings, and EQ score) and its interaction with tone trial type. Then, we compared the models with the basic model.

3. Results

3.1. Neurophysiological sensory attenuation

The average ERPs of each condition are depicted in Fig. 2B. The estimated basic model for each component is summarised in Table 1. For the N1, the decomposition of fixed-effects contributions by using the anova function revealed that the effect of tone trial type was significant, $F(3144) = 4.37$, $p < .001$, $\eta_p^2 = 0.08$. The N1 amplitudes of self-generated ($M = -6.26$ µV, $p = .076$) and other-generated ($M = -6.35$ µV, $p = .14$) were only descriptively reduced, but the amplitudes for cued tones ($M = -5.87$ µV, $p = .004$) were significantly reduced compared to those of the un-cued tones ($M = -7.28$ µV; see Fig. 2C). The remaining pairwise comparisons were not significant (all $p > .99$).

For the P2, the effect of tone trial type was significant, $F(3138) = 58.12$, $p < .001$, $\eta_p^2 = 0.56$. The P2 amplitudes of self-generated, other-generated, and cued tones (all $p < .001$) were significantly reduced compared to those of the un-cued tones ($M = 8.94$ µV; see Fig. 2 C). Furthermore, there were significantly reduced amplitudes for self-generated ($M = 4.52$ µV, $p < .001$) and for other-generated tones ($M = 3.72$ µV, $p < .001$) in comparison to cued tones ($M = 6.59$ µV). The amplitudes between the self- and other-generated tones did not differ significantly ($p = .41$).

3.2. Effect of individualism

For both individualism ($M = 3.74$, $SD = 0.64$, range: 2.60–5.27) and collectivism ($M = 3.60$, $SD = 0.79$, range: 1.40–5.50) the sample mean was around the medium score of the scale (3.5) and spanned over the possible range (i.e., 1–6). Individualism and collectivism mean scores did not differ significantly, $t(48) = 1.121$, $p = .27$, $d = 0.160$. Indeed, the difference values (individualism-collectivism) were normally distributed around 0.14 ($SD = 0.87$, range: −2.53 to 2.37, Shapiro-Wilk: $p = .50$). On the individualism sub-scores, participants scored higher descriptively on responsibility ($M = 4.49$, $SD = 0.63$, range: 3.00–6.00) and uniqueness ($M = 4.18$, $SD = 0.95$, range: 2.50–6.00) than competitiveness ($M = 3.06$, $SD = 0.87$, range: 1.57–5.57). On the collectivism sub-scores, participants scored higher descriptively on advice ($M = 3.73$, $SD = 1.03$, range: 1.29–5.86) than on harmony ($M = 3.30$, $SD = 0.90$, range: 1.50–5.75). The calculated models including culture-related variables and the respective comparisons are reported in Table 2 for the N1 and in Table 3 for the P2 (also for anomalous perception and empathy models). For both N1 and P2, none of the models including a questionnaire (sub)score provided a significantly better fit to the data than the basic model.

4. Discussion

In this study, we compared N1 and P2 amplitudes of self-generated, other-generated, cued and un-cued tones to investigate neurophysiological SA, with a particular focus on the other-generated condition and

Table 1
Summary of the Results of the Basic Model for the N1 and P2 Component.

	Fixed Effects								R^2	
	TTT	Est.	*SE*	*t*	*p*	ICC	AIC	BIC	Cond.	Marg.
N1	Un-cued	-7.28	0.49	-14.80	< .001	0.66	943.0	962.7	0.67	0.02
	Cued	1.41	0.40	3.49	< .001					
	Other	0.92	0.40	2.28	.024					
	Self	1.02	0.40	2.52	.013					
P2	Un-cued	8.95	0.49	18.30	< .001	0.61	913.5	932.9	0.71	0.27
	Cued	-2.36	0.43	-5.43	< .001					
	Other	-5.22	0.43	-12.05	< .001					
	Self	-4.43	0.43	-10.21	< .001					

Note. The basic model for N1 (49 participants) and P2 (47 participants) amplitudes: Amplitude ~ tone trial type + (1|participant). TTT = tone trial type, Est. = Estimate, Cond. = Conditional, Marg. = Marginal

Table 2
Model Comparisons for the N1 Component.

Basic model: N1 amplitude ~ tone trial type + (1|participant)

AIC	BIC	LL	df
942.5	962.2	-465.2	6

Further models: N1 amplitude ~ tone trial type * variable + (1|participant)

Effect	Variable	AIC	BIC	LL	df	LRT $\chi^2(4)$	*P*	BF
Culture (AICS)	Individualism	949.2	982.0	-464.6	10	1.247	.87	4.86e-5
	Competitiveness	948.0	980.8	-464.0	10	2.448	.65	8.85e-5
	Responsibility	944.6	977.4	-462.3	10	5.890	.21	4.95e-4
	Uniqueness	947.0	979.7	-463.5	10	3.528	.47	1.52e-4
	Collectivism	947.0	979.8	-463.5	10	3.501	.48	1.50e-4
	Advice	945.5	978.3	-462.7	10	5.014	.29	3.19e-4
	Harmony	948.5	981.3	-464.2	10	2.010	.73	7.11e-5
Anomalous Perception (CAPS)	Total score	949.0	981.8	-464.5	10	1.454	.83	5.39e-5
	Distress	948.8	981.6	-464.4	10	1.706	.79	6.11e-5
	Frequency	947.7	980.5	-463.8	10	2.797	.59	1.05e-4
	Intrusiveness	948.5	981.3	-464.3	10	1.971	.74	6.97e-5
Empathy (EQ)		948.3	981.0	-464.1	10	2.233	.69	7.95e-5

Note. For all models the data of 49 participants was used (total of 196 observations). AIC = Aikake information criterion, BIC = Bayesian information criterion, LL = loglikelihood, LRT = Likelihood ratio test, BF = Bayes Factor.

Table 3
Model Comparisons for the P2 Component.

Basic model: P2 amplitude ~ tone trial type + (1|participant)

AIC	BIC	LL	df
980.31	999.98	-484.16	6

Further models: P2 amplitude ~ tone trial type * variable + (1|participant)

Effect	Variable	AIC	BIC	LL	df	LRT $\chi^2(4)$	*P*	BF
Culture (AICS)	Individualism	985.4	1018.2	-482.7	10	2.912	.57	1.12e-4
	Competitiveness	984.3	1017.1	-482.2	10	3.979	.41	1.90e-4
	Responsibility	983.8	1016.6	-481.9	10	4.502	.34	2.47e-4
	Uniqueness	987.2	1019.9	-483.6	10	1.155	.89	4.64e-5
	Collectivism	987.3	1020.1	-483.7	10	0.973	.91	4.23e-5
	Advice	987.9	1020.6	-483.9	10	0.460	.98	3.28e-5
	Harmony	987.1	1019.9	-483.5	10	1.221	.87	4.79e-5
Anomalous Perception (CAPS)	Total score	986.9	1019.6	-483.4	10	1.460	.83	5.40e-5
	Distress	985.7	1018.5	-482.9	10	2.578	.63	9.45e-5
	Frequency	985.5	1018.3	-482.8	10	2.807	.59	1.06e-4
	Intrusiveness	985.6	1018.4	-482.8	10	2.692	.61	1.00e-4
Empathy (EQ)		987.6	1020.4	-483.8	10	0.734	.95	3.76e-5

Note. For all models the data of 49 participants was used (total of 196 observations). AIC = Aikake information criterion, BIC = Bayesian information criterion, LL = loglikelihood, LRT = Likelihood ratio test, BF = Bayes Factor.

thus action observation. The other-generation was implemented as a first-person perspective on the button presses of the other participant in a simultaneous EEG acquisition. We added an action-outcome delay of 350 ms throughout to allow the prediction of the timing of tone onset also for the observed button press and the visual cue. Our results do not show clear neurophysiological SA for the N1, as the amplitudes for the self and other condition were only descriptively reduced relative to un-cued tones and not reduced relative to cued tones. We found a significant N1 amplitude reduction for cued external tones compared to un-cued ones. For the P2 component a different pattern emerged. We found a significant attenuation for self- and other-generated as well as for cued external tones compared to un-cued tones, with this attenuation being stronger for both self- and other-generated tones relative to cued tones. The strength of P2 attenuation for self- and other-generated tones did not differ. For both N1 and P2 components, the models including questionnaire scores related to cultural differences did not provide a significantly better fit than the basic model with only the tone trial type as fixed effect. Thus, our results indicated no effects of individualism or collectivism on neurophysiological SA. The models including the control variables anomalous perception and empathy were not significantly better either, thus ruling out potential confounding effects of anomalous perception or empathy.

4.1. *Neurophysiological sensory attenuation in action observation*

4.1.1. *N1 component*

The neurophysiological SA in N1 was significant for cued (relative to un-cued) tones only. Although the amplitudes for cued tones did not differ from those of self- and other-generated tones, and the latter were descriptively reduced relative to un-cued external tones, we did not find significant evidence of SA in the N1 for these conditions. A possible interpretation of the overall pattern of results is that temporal predictability fully accounts for the (descriptive) N1 attenuation for self, other, and cued tones. This influence of temporal predictability on N1 attenuation is in line with previous studies limited to action performance (e.g., Harrison et al., 2021; Kaiser & Schütz-Bosbach, 2018), and the interpretation that SA is based on general predictive mechanisms as proposed in recent reviews (Dogge et al., 2019; Korka et al., 2022).

However, the fact that we did not observe neurophysiological SA in the N1 for self-generation relative to un-cued external tones should be carefully discussed as this effect has been consistently reported in a wide range of studies (Horváth, 2015; Korka et al., 2022), including studies with the same or longer action-outcome delays (Klaffehn et al., 2019; Lange, 2011), and a lack of attenuation is rare (Ody et al., 2022). The paradigm applied in this study was similar to those applied in previous studies (e.g., Ghio et al., 2021), and the only methodological change that might have influenced the attenuation effect was the use of a mirror construction to reflect the button press in first-person perspective. Participants did not report any problems getting accustomed to the unusual view on their actions and we assumed this would not impact the attenuation effect. This modification, however, might have had an impact on the prediction mechanisms especially for self-generated tones, which should be further investigated in future research.

It should be noted that there are previous studies that investigated the effect of temporal predictability on neurophysiological SA, albeit limited to action performance, and did not support the interpretation in terms of general predictive mechanisms (Klaffehn et al., 2019; Lange, 2011). Possibly, the predictability as provided by the cue in these studies was not as high as the predictability the own action affords (Kiepe et al., 2021).

Still another interpretation of neurophysiological SA is in terms of attention as it has been shown that tones that are more attended than others elicit larger N1 amplitudes (Horváth, 2015; Saupe et al., 2013). As all conditions but the un-cued condition had an event preceding the tone that could have drawn attention away from the tone itself, perhaps more so for the cue than the (observed) action, we cannot exclude attention effects as a possible alternative explanation for the N1 results. Yet, Timm et al. (2013) and Harrison et al. (2021) found no N1 modulation by (visual) attention and studies that included external tones in the same experimental block to account for effects of attention reported N1 attenuation for self-generated tones (Ghio et al., 2017; Knolle et al., 2013), though this was not the case for other-generated tones (Ghio et al., 2017).

In conclusion, our pattern of results seems to speak in favour of general predictive mechanisms with similar N1 attenuation for self-, other-generated and cued tones, at least when the event-outcome delay is sufficient to enable prediction. However, a firm conclusion cannot be drawn from this study: General predictive mechanisms as applied for cued external and possibly for other-generated tones may simply result in similar N1 attenuation as motor- or intention-based mechanisms, and we still cannot exclude potential interactions between motor or intention-based and general mechanisms (Korka et al., 2022).

4.1.2. P2 component

We found an attenuation of the P2 amplitude for self-, other-generated and cued external tones compared to un-cued tones. Interestingly, we additionally found that amplitudes of self- and other-generated tones were attenuated also compared to cued external tones. While an interpretation in terms of motor-based mechanisms cannot explain the P2 attenuation for cued tones, a general effect of temporal predictability seems plausible to explain the P2 attenuation for this condition and, possibly in part, also for the other conditions, for which an attenuation was found. Moreover, action performance and even action observation may be better predictors than visual cues that were matched to button presses (and not to finger movements) resulting in a stronger attenuation. In line with this, van Laarhoven et al. (2020) showed that the P2 was attenuated for sounds of observed handclaps compared to sounds without the visual input and this attenuation was reduced, though still significant, when the timing was not predictable. On the other hand, Klaffehn et al. (2019) reported P2 attenuation only for self-generated tones that were not delayed, while P2 amplitudes for delayed self-generated tones (similar to the self-generated tones in our study) were not attenuated at all. Further, Harrison et al. (2021) found that P2 amplitudes of cued external and cued self-generated tones (i.e., participants had no temporal control) were not attenuated; only the amplitudes of temporally controlled self-generated tones were. The P2 attenuation that we observed may therefore not relate to the temporal predictability of the tone but to other factors, which is also in line with our findings for the N1 amplitude.

Another interpretation of the P2 is in terms of the agency judgement (Knolle et al., 2013; Seidel et al., 2021; Timm et al., 2016). Judgment or belief of agency was also found to influence perceptual SA, as self-generated tones were attenuated compared to tones that were self-generated but believed to be other-generated (Desantis et al., 2012). A recent finding, that potentially supports this interpretation, is that P2 attenuation was found for tones produced by voluntarily elicited involuntary actions that included a finger-on-finger press but not those that were achieved via an electrical stimulator (Jack et al., 2021). Although, in both cases the participants initiated the action, the electrical stimulator as an external intermediary device might have led participants to shift the agency to it rather than attribute the agency to themselves. In our study, the agency for self- and other-generated tones could be clearly attributed to oneself or the observed person. Yet, the P2 attenuation was not significantly different in these conditions. Hence, distinguishing between one's own and other's action-consequences based on the P2 amplitudes alone was not possible. In line with this, the observer P2 attenuation was found in both our previous studies on neurophysiological SA in action observation (Ghio et al., 2021; Ghio et al., 2017), albeit weaker than for self-generation in Ghio et al. (2017). Interpreting the P2 component as a reflection of the agency judgement, this would mean that the tones were judged to have been generated by the observed person similarly to the tones being judged to have been generated by oneself and so the P2 attenuation would reflect the attribution to a human agent. As a live observation (i.e., simultaneous EEG acquisition) was done in the present study and Ghio et al. (2021), this setting might determine the strong observer P2 attenuation, regardless of the perspective which differed between the studies (Ghio et al., 2021; Ghio et al., 2017).

Regarding the P2 attenuation for cued tones compared to un-cued tones, it should be noted that the cues in this study were white circles depicted on top of the button that the participants pressed to elicit the tones. Because of the temporal, spatial, and outcome similarity, participants may have simulated the action or regarded the cue as a type of agent, though not as strongly as for another human being. Non-human agents can be attributed agency and intention in a similar manner as human agents (i.e., anthropomorphism) in case of robotic agents that mimic the human appearance (Epley et al., 2007): Indeed, studies on action observation suggested that the observation of robots might become comparable to the observation of other humans, when perceived more frequently (Press, 2011). However, the cue in the present study was not human-like and was presented on the computer screen, which could have diminished its interpretation as an agent. P2 attenuation thus might reflect the attribution of the tone to any agent (oneself, other human or human-like external) and not be essential to the self-external distinction of event-outcomes.

4.2. Individualism and other potential effects on sensory attenuation

Our results on the neurophysiological SA in observation, at least for the P2 component, are in line with the behavioural findings of perceptual SA in observers in the study by Sato (2008) but not with the findings of Weiss et al. (2011). Both studies compared self- and other-generated to cued external tones but only Sato (2008) found perceptual SA for both self- and other-generated tones, while Weiss et al. (2011) found a significantly lower perceived loudness for self- relative to other-generated tones. The divergence of the two behavioural studies in perceptual SA for observation was proposed to be a cultural effect (Cao & Gross, 2015). At a cultural level, western countries, such as Germany, tend to score high on individualism (Hofstede, 2011). According to Cao and Gross (2015), the degree of individualism served as possible explanation of the lack of perceptual SA for other-generated tones in their British sample and in the German sample of Weiss et al. (2011). Yet, our sample in this study consisted of mostly German natives and (possibly assimilated) migrants in Germany. With the AICS we measured individualism and collectivism in two separate scales, including various subscales (Shulruf et al., 2011; Shulruf et al., 2007). Our sample was not more individualistic than it was collectivistic and many participants scored similarly on both scores. Hence, we had few participants that could be clearly defined as individualistic or clearly defined as collectivistic. This is in line with findings that place cultures of Western European countries, like Germany, in the middle instead of at the individualistic end of the continuum (Kitayama et al., 2009). Importantly, no model including scores related to individualism (or collectivism) explained the ERP data better than the basic model including only tone trial type as predictor, which can be interpreted in different ways. On the one hand, our finding might indicate that there is no effect of individualism (or culture in general) on neurophysiological SA. This does, however, not exclude that perceptual SA is affected by individualism (Cao & Gross, 2015). For other modalities it has been suggested that perceptual and neurophysiological SA are not equivalent (visual modality: Roussel et al., 2014) and are the result of different mechanisms (tactile modality: Palmer et al., 2016). On the other hand, our negative finding should be interpreted with caution as the variability of individualism (and collectivism) scores in our sample may not have been large enough or the statistical power of our study may have not sufficed to detect a relationship between individualism and neurophysiological SA.

Similar to the results for the culture-related scores, we did not find effects of empathy and anomalous perception on N1 and P2 amplitudes. Especially with respect to the latter, an effect would have been conceivable, as in previous work SA has been suggested to be related to delusional thinking and hallucinations. For example, diminished N1 reductions have been found in schizophrenic patients (Ford et al., 2007) and linked to the severity of hallucinations (Heinks-Maldonado et al., 2007). In a sample of healthy participants, Cao and Gross (2015) found that the difference values representing perceptual SA in action performance were related to delusional ideation scores (in the British sample). Given that our sample consisted entirely of people stating to have no psychiatric disorders and that we applied a questionnaire focusing on anomalous perception (Bell et al., 2006) instead of more general delusional ideation, the effect might have been too weak to detect. Our negative finding in this respect can thus be interpreted to be exploratory and needs to be tested further.

4.3. Limitations and outlook

It should be noted that the motor correction we applied has been criticised, for actions with no apparent outcome and actions that elicit a tone do not have the same action-related activity (Horváth, 2015). A recent study found that such action-related differences might affect the N1 and P2, in that the N1 was more attenuated and the P2 less attenuated as a result of the correction (Neszmélyi et al., 2022). Because of this and because we found a significant difference in interval length for the other-generation blocks (observer auditory-motor and motor-only) our results have to be interpreted with caution.

A live observation could be affected by other processes than an observation via video in a single EEG acquisition, like in Ghio et al. (2017), where the P2 attenuation for other-generated tones was weaker than for self-generated tones. We further did not account for factors specific to this setting, such as degree of acquaintance or likeability of the paired participant, which might influence SA in observation.

5. Conclusion

The present study extends previous findings on the neurophysiological SA effect for observed actions. With a first-person perspective on the observed action and a temporal delay between button press and tone onset, other-generated tones led to similar N1 amplitudes compared to self-generated and cued external tones, while a significant N1 attenuation was only found for cued relative to un-cued tones. This pattern of results might indicate that N1 attenuation depends mainly on the temporal predictability of the tone and therefore reflects more general predictive mechanisms (Dogge et al., 2019; Korka et al., 2022), though attentional effects cannot be excluded (Horváth, 2015). Further, the P2 was also comparable between self- and other-generated tones, and there was a significant attenuation relative to cued and un-cued tones. This attenuation could partly depend on temporal predictability, as the amplitudes of cued external tones were also attenuated relative to un-cued ones. The stronger P2 attenuation for self- and other-generated tones could indicate the clearer agency attribution to oneself or another person, which was not comparable in the cued condition. Importantly, with our sample we did not find an effect of individualism, collectivism, anomalous perception, or empathy on the attenuation of these components.

Author Note

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.biopsycho.2023.108575.

References

Arimie, C. O., Biu, E. O., & Ijomah, M. A. (2020). Outlier detection and effects on modeling. *OPen Access Library Journal, 07*(09), 1–30. https://doi.org/10.4236/oalib.1106619

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). The Cardiff Anomalous Perceptions Scale (CAPS): a new validated measure of anomalous perceptual experience. *Schizophrenia Bulletin, 32*(2), 366–377. https://doi.org/10.1093/schbul/sbj014

Bellebaum, C., Ghio, M., Wollmer, M., Weismuller, B., & Thoma, P. (2020). The role of trait empathy in the processing of observed actions in a false-belief task. *Social Cognitive and Affective Neuroscience, 15*(1), 53–61. https://doi.org/10.1093/scan/nsaa009

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: estimation of effect size indices and standatdized parameters. *Journal of Open Source Software, 5*(56), 2815. https://doi.org/10.21105/joss.02815

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *NeuroReport, 12*(9), 1879–1884.

Blakemore, S.-J., & Sirigu, A. (2003). Action prediction in the cerebellum and in the parietal lobe. *Exp Brain Res, 153*(2), 239–245. https://doi.org/10.1007/s00221-003-1597-z

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2013). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Social Neuroscience* (pp. 133–141). Psychology Press,.

Cao, L., & Gross, J. (2015). Cultural differences in perceiving sounds generated by others: self matters. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01865

Cao, L., Veniero, D., Thut, G., & Gross, J. (2017). Role of the cerebellum in adaptation to delayed action effects. *Current Biology, 27*(16), 2442–2451 e2443. https://doi.org/10.1016/j.cub.2017.06.074

Cook, R. D. (2000). Detection of influential observation in linear regression. *Technometrics, 42*(1), 65–68. https://doi.org/10.1080/00401706.2000.10485981

Desantis, A., Weiss, C., Schutz-Bosbach, S., & Waszak, F. (2012). Believing and perceiving: authorship belief modulates sensory attenuation. *PLoS ONE, 7*(5), Article e37959. https://doi.org/10.1371/journal.pone.0037959

Dogge, M., Custers, R., & Aarts, H. (2019). Moving forward: on the limits of motor-based forward models. *Trends in Cognitive Sciences, 23*(9), 743–753. https://doi.org/10.1016/j.tics.2019.06.008

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., & Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology, 44*(4), 522–529. https://doi.org/10.1111/j.1469-8986.2007.00533.x

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological Evidence of Corollary Discharge Dysfunction in Schizophrenia. *American Journal of Psychiatry, 158*(12), 2069–2071. https://doi.org/10.1176/appi.ajp.158.12.2069

Ghio, M., Egan, S., & Bellebaum, C. (2021). Similarities and differences between performers and observers in processing auditory action consequences: Evidence from simultaneous EEG acquisition. *Journal of Cognitive Neuroscience, 33*(4), 683–694. https://doi.org/10.1162/jocn_a_01671

Ghio, M., Scharmach, K., & Bellebaum, C. (2017). ERP correlates of processing the auditory consequences of own versus observed actions. *Psychophysiology, 55*(6), Article e13048. https://doi.org/10.1111/psyp.13048

Harrison, A. W., Mannion, D. J., Jack, B. N., Griffiths, O., Hughes, G., & Whitford, T. J. (2021). Sensory attenuation is modulated by the contrasting effects of predictability and control. *NeuroImage, 237*, Article 118103. https://doi.org/10.1016/j.neuroimage.2021.118103

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Archives of General Psychiatry, 64*(3), 286–296. https://doi.org/10.1001/archpsyc.64.3.286

Hofstede, G. (2011). Dimensionalizing cultures: the hofstede model in context. *Online readings in Psychology and Culture, 2*(1). https://doi.org/10.9707/2307-0919.1014

Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research, 1626*, 54–65. https://doi.org/10.1016/j.brainres.2015.03.038

Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action–sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience, 24*(9), 1919–1931. https://doi.org/10.1162/jocn_a_00215

Hughes, G., Desantis, A., & Waszak, F. (2013). 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. https://doi.org/10.1037/a0028566

Jack, B. N., Chilver, M. R., Vickery, R. M., Birznieks, I., Krstanoska-Blazeska, K., Whitford, T. J., & Griffiths, O. (2021). Movement planning determines sensory suppression: an event-related potential study. *Journal of Cognitive Neuroscience, 33*(12), 2427–2439. https://doi.org/10.1162/jocn_a_01747

Kaiser, J., & Schütz-Bosbach, S. (2018). Sensory attenuation of self-produced signals does not rely on self-specific motor predictions. *European Journal of Neuroscience, 47*(11), 1303–1310. https://doi.org/10.1111/ejn.13931

Kiepe, F., Kraus, N., & Hesselmann, G. (2021). Sensory attenuation in the auditory modality as a window into predictive processing. *Frontiers in Human Neuroscience, 15*. https://doi.org/10.3389/fnhum.2021.704668

Kitayama, S., Park, H., Sevincer, A. T., Karasawa, M., & Uskul, A. K. (2009). A cultural task analysis of implicit independence: comparing North America, Western Europe, and East Asia. *Journal of Personality and Social Psychology, 97*(2), 236–255. https://doi.org/10.1037/a0015999

Klaffehn, A. L., Baess, P., Kunde, W., & Pfister, R. (2019). Sensory attenuation prevails when controlling for temporal predictability of self-and externally generated tones. *Neuropsychologia, 132*, Article 107145. https://doi.org/10.1016/j.neuropsychologia.2019.107145

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex, 49*(9), 2449–2461. https://doi.org/10.1016/j.cortex.2012.12.012

Korka, B., Widmann, A., Waszak, F., Darriba, Á., & Schröger, E. (2022). The auditory brain in action: Intention determines predictive processing in the auditory system—A review of current paradigms and findings. *Psychonomic Bulletin & Review, 29*(2), 321–342. https://doi.org/10.3758/s13423-021-01992-z

Krugwasser, A. R., Stern, Y., Faivre, N., Harel, E. V., & Salomon, R. (2022). Impaired sense of agency and associated confidence in psychosis. *Schizophrenia, 8*(1), 32. https://doi.org/10.1038/s41537-022-00212-4

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lange, K. (2011). The reduced N1 to self-generated tones: an effect of temporal predictability? *Psychophysiology, 48*(8), 1088–1095. https://doi.org/10.1111/j.1469-8986.2010.01174.x

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the Empathy Quotient. *Psychological Medicine, 34*(5), 911–920. https://doi.org/10.1017/S0033291703001624

Lenth, R.V., 2022, emmeans: Estimated Marginal Means, aka Least-Squares Means. In ⟨https://CRAN.R-project.org/package=emmeans⟩.

Luck, S. J., & Gaspelin, N. (2017). How to Get Statistically Significant Effects in Any ERP Experiment (and Why You Shouldn't). *Psychophysiology, 54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139

McCarthy, G., & Donchin, E. (1976). The effects of temporal and event uncertainty in determining the waveforms of the auditory event related potential (ERP). *Psychophysiology, 13*(6), 581–590. https://doi.org/10.1111/j.1469-8986.1976.tb00885.x

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*. https://doi.org/10.1016/j.jml.2020.104092

Neszmélyi, B., & Horváth, J. (2017). Consequences matter: Self-induced tones are used as feedback to optimize tone-eliciting actions. *Psychophysiology, 54*(6), 904–915. https://doi.org/10.1111/psyp.12845

Neszmélyi, B., Nevelo, D., & Horváth, J. (2022). Turning a blind eye to motor differences leads to bias in estimating action-related auditory ERP attenuation. *Biological Psychology, 173*, Article 108387. https://doi.org/10.1016/j.biopsycho.2022.108387

Nieuwenhuis, R., Grotenhuis, M. T., & Pelzer, B. (2012). influence.ME: tools for detecting influential data in mixed effects models. *R Journal, 4*(2), 38–47.

Ody, E., Straube, B., He, Y., & Kircher, T. (2022). Perception of self-and externally-generated visual stimuli: Evidence from EEG and behaviour, 2022.2005.2030.492531 *bioRxiv*. https://doi.org/10.1101/2022.05.30.492531.

Palmer, C. E., Davare, M., & Kilner, J. M. (2016). Physiological and perceptual sensory attenuation have different underlying neurophysiological correlates. *Journal of Neuroscience, 36*(42), 10803–10812. https://doi.org/10.1523/JNEUROSCI.1694-16.2016

Pinheiro, A. P., Schwartze, M., Gutierrez, F., & Kotz, S. A. (2019). When temporal prediction errs: ERP responses to delayed action-feedback onset. *Neuropsychologia, 134*, Article 107200. https://doi.org/10.1016/j.neuropsychologia.2019.107200

Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. *Neurosci Biobehav Rev, 35*(6), 1410–1418. https://doi.org/10.1016/j.neubiorev.2011.03.004

R. Core Team, 2022, R: A language and environment for statistical computing. In R Foundation for Statistical Computing, Vienna, Austria. ⟨http://www.R-project.org/⟩.

Roussel, C., Hughes, G., & Waszak, F. (2014). Action prediction modulates both neurophysiological and psychophysical indices of sensory attenuation. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00115

Sato, A. (2008). Action observation modulates auditory perception of the consequence of others' actions. *Consciousness and cognition, 17*(4), 1219–1227. https://doi.org/10.1016/j.concog.2008.01.003

Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial suppression of self-generated sounds and its dependence on attention. *International Journal of Psychophysiology, 90*(3), 300–310. https://doi.org/10.1016/j.ijpsycho.2013.09.006

Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain Responses. *Science, 181*(4095), 175–177. https://doi.org/10.1126/science.181.4095.175

Seidel, A., Ghio, M., Studer, B., & Bellebaum, C. (2021). Illusion of control affects ERP amplitude reductions for auditory outcomes of self-generated actions. *Psychophysiology, 58*(5), Article e13792. https://doi.org/10.1111/psyp.13792

Seidel, A., Weber, C., Ghio, M., & Bellebaum, C. (2023). My view on your actions: Dynamic changes in viewpoint-dependent auditory ERP attenuation during action observation. *Cognitive, Affective, & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-023-01083-7

Shulruf, B., Alesi, M., Ciochină, L., Faria, L., Hattie, J., Hong, F., Pepi, A.-M., & Watkins, D. (2011). Measuring collectivism and individualism in the third millennium. *Social Behavior and Personality: an International Journal, 39*(2), 173–187. https://doi.org/10.2224/sbp.2011.39.2.173

Shulruf, B., Hattie, J., & Dixon, R. (2007). Development of a New Measurement Tool for Individualism and Collectivism. *Journal of Psychoeducational Assessment, 25*(4), 385–401. https://doi.org/10.1177/0734282906298992

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. https://doi.org/10.1007/s00221-012-3204-7

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*(7), 1481–1489. https://doi.org/10.1162/jocn_a_00552

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), 1–11. https://doi.org/10.1186/1471-2202-14-2

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. https://doi.org/10.1016/j.cortex.2016.03.018

van Laarhoven, T., Stekelenburg, J. J., & Vroomen, J. (2020). Suppression of the auditory N1 by visual anticipatory motion is modulated by temporal and identity predictability. *Psychophysiology, 58*(3). https://doi.org/10.1111/psyp.13749

Volpert-Esmond, H. I., Page-Gould, E., & Bartholow, B. D. (2021). Using multilevel models for the analysis of event-related potentials. *International Journal of Psychophysiology, 162*, 145–156. https://doi.org/10.1016/j.ijpsycho.2021.02.006

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011). The self in action effects: Selective attenuation of self-generated sounds. *Cognition, 121*(2), 207–218. https://doi.org/10.1016/j.cognition.2011.06.011

Whitford, T. J. (2019). Speaking-induced suppression of the auditory cortex in humans and its relevance to schizophrenia. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4*(9), 791–804. https://doi.org/10.1016/j.bpsc.2019.05.011

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag,. ⟨https://ggplot2.tidyverse.org⟩.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1431), 593–602. https://doi.org/10.1098/rstb.2002.1238

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology, 11*(18), R729–R732.

Yamada, A.-M., & Singelis, T. M. (1999). Biculturalism and self-construal. *International Journal of Intercultural Relations, 23*(5), 697–709. https://doi.org/10.1016/S0147-1767(99)00016-4

6.2 Study 2

I, You, Robot: Attenuation for Auditory Outcomes of Actions Performed by Different Agents Shows Distinct Patterns for N1 and P2 Amplitudes

Sophie Egan, Constanze Weber, Marta Ghio, and Christian Bellebaum

Institute of Experimental Psychology, Faculty of Mathematics and Natural Sciences, Heinrich Heine University Düsseldorf

Author Note

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Email: Sophie.Egan@hhu.de, Constanze.Weber@hhu.de, Marta.Ghio@hhu.de, Christian.Bellebaum@hhu.de

Postal address: Universitätsstraße 1, 40225 Düsseldorf, Germany

Corresponding author: Sophie Egan

Word count: 8493

Abstract word count: 243

Abstract

Compared to external tones, previous studies found reduced N1 and P2 amplitudes in event-related potentials not only for self-generated tones, referred to as sensory attenuation, but also for tones generated by a person the participant observed. The P2 in this situation was attenuated, even when accounting for temporal predictability. P2 attenuation has also been related to the sense of agency. To test whether the P2 attenuation in action observation depends on the attribution of agency to a human agent, we added conditions with robotic hands performing the tone-eliciting action. Therefore, we compared N1 and P2 amplitudes of tones generated by a self-performed action of the participant, by an observed action of another person using either their own or a robotic hand, and by the computer shown as a robotic hand, with those of conditions with visual cues predicting the tones and with unpredictable tones. Only N1 amplitudes in action observation with a human agent (human or robotic hand) were significantly attenuated compared to unpredictable external tones. The P2 was attenuated for the self-generation, action-observation and cue conditions compared to unpredictable external tones, but the P2 of self-generated tones was further attenuated compared to all other conditions. Thus, while the N1 may be affected by (observed) human agency, the P2 showed attenuation effects related to temporal predictability and self-agency. This contrasts with previous findings in action observation suggesting also other-agency effects. Future research could further examine the modulation of attenuation by the action-observation setting.

Keywords: auditory sensory attenuation, action observation, temporal predictability, sense of agency, EEG

I, You, Robot: Attenuation for Auditory Outcomes of Actions Performed by Different Agents Shows Distinct Patterns for N1 and P2 Amplitudes

Sensory attenuation (SA) was defined as a reduction in the perceptual or neurophysiological response to stimuli that are self-generated compared to external stimuli that are otherwise identical (for reviews see Dogge et al., 2019; Hughes et al., 2013; Kiepe et al., 2021). For the auditory modality that means perceiving self-generated tones as less loud and having reduced neurophysiological responses to self-generated tones compared to identical external tones. The latter is reflected in reduced amplitudes of the event-related potential (ERP) components N1 and P2, which are typically analysed in studies on auditory SA applying electroencephalography (EEG; Horváth, 2015; Kiepe et al., 2021).

SA has traditionally been explained by forward model predictions of the sensory input based on the efference copy of the motor command (Wolpert & Flanagan, 2001). Because this would make SA specific to self-generation, SA has been considered as an implicit measure of the sense of agency (Hughes et al., 2013; Kiepe et al., 2021). Indeed, neurophysiological SA (N1, P2, and N1-P2 complex) was found to occur only for voluntary as opposed to involuntary action (Jack et al., 2021; Timm et al., 2014). In the forward model framework, the cerebellum has been identified as a relevant brain structure for both motor-to-somatosensory (e.g., Blakemore et al., 2001) and motor-to-auditory predictions: For neurophysiological SA in the auditory modality, N1 attenuation was absent in cerebellar patients but P2 attenuation was intact (Knolle et al., 2013), suggesting a functional dissociation between the two components, with only the N1 reflecting forward model predictions. However, forward models do not suffice to explain all SA-related phenomena reported in the literature, such as attenuation of N1 and P2 without a contingent action-effect relation (e.g., Horváth et al., 2012), N1 attenuation without action (e.g., Harrison et al., 2021), and the typically rapid learning of arbitrary action-effect associations (Dogge et al., 2019). Instead, it has been suggested that a combination of different mechanisms produces neurophysiological SA (Dogge et al., 2019; Horváth, 2015; Hughes et al., 2013). According to accounts claiming that SA can be explained by general predictive mechanisms, a similar attenuation of neurophysiological activity might occur for predictable external tones relative to unpredictable ones as for self-generated tones (Dogge et al., 2019; Kiepe et al., 2021; Korka et al., 2022). Findings regarding this are mixed, but some studies have indeed found such attenuation for the N1 (e.g., Harrison et al., 2021) and the P2 (e.g., Schafer & Marcus, 1973).

If SA underlies the sense of agency (Hughes et al., 2013; Kiepe et al., 2021), uncertainties regarding the agency would arise when there is no SA relative to externally generated tones (cf. Sato, 2008). In line with this, in patients with schizophrenia, which has been related to impaired agency attribution (Franck et al., 2001), N1 attenuation was reduced compared to healthy controls (Ford & Mathalon, 2012). Studies further found significant correlations between hallucination severity as well as agency misattribution and N1 attenuation in patients with schizophrenia (Heinks-Maldonado et al., 2007), and between delusional ideation in healthy participants and perceptual SA (Cao & Gross, 2015). However, Egan et al. (2023) found no significant relation between anomalous perception in healthy participants and neurophysiological SA (i.e., N1 or P2 attenuation). Experimental manipulations of agency with healthy participants have shown diverging results. On the one hand, agency but not neurophysiological SA (i.e., N1 or P2 attenuation) was affected by filling the delay between action and effect (Weller et al., 2017). On the other hand, previous studies imply a relation of agency judgement with the P2 component (Seidel et al., 2021; Timm et al., 2016), and perceptual SA (Desantis et al., 2012).

Action observation is similar to self-generation not only in terms of attribution to an agent but also to some extent in brain activation (for reviews regarding the mirror neuron system see Ferrari & Rizzolatti, 2014; Iacoboni, 2009). In the forward model framework, it was proposed that the mirror neuron system, and the action simulation it might entail, enables motor-based predictions also for observed action-outcomes, similar to the process for self-generation (Wolpert et al., 2003). As such, SA might be extended to other-generated tones. However, the observed action makes the tone temporally predictable, which on its own might lead to attenuated N1 and P2 amplitudes (Hughes et al., 2013).

Previous SA studies in action observation provided mixed results. Perceptual SA was reported to be both present (Sato, 2008) and absent (Weiss et al., 2011a) in studies applying similar experimental paradigms, with a potential modulation by the participants' individualism (i.e., a cultural dimension; Cao & Gross, 2015). Concerning neurophysiological SA, N1 attenuation for tones elicited by observed actions was only found with sufficient delay between action and outcome, making the tones temporally predictable (Ghio et al., 2021; Ghio et al., 2018). Indeed, recent studies found no amplitude difference between cued tones and tones generated by observed actions (Egan et al., 2023); and between cued and self-generated tones (Egan et al., 2023; Harrison et al., 2021), further suggesting that N1 attenuation reflects temporal predictability. The P2, on the other hand, was consistently attenuated in action observation (Egan et al., 2023; Ghio et al., 2021) and, as shown by a recent study, the attenuation was stronger than for tones that were visually cued without action observation (Egan et al., 2023). The same study also showed that the P2 for cued tones was attenuated relative to un-cued, and thus unpredictable, tones, suggesting that the P2 might reflect in part predictability and in part agency attribution. However, the cue and the observation conditions in this study (Egan et al., 2023) differed not only in the attribution of the tone to a human agent, but also in the action-effect association, which was lacking in the cue condition.

In the present study, we further investigated whether neurophysiological SA, represented by the N1 and particularly the P2, reflects agency attribution, with a focus on action observation. We did this by extending the previously applied paradigm (Egan et al., 2023) with two additional action-observation conditions, in which the button press was performed by a robotic hand: In one condition the robotic hand was said to be controlled by another person and in the other by the computer. Unlike the cue condition, both conditions included an action-effect association. For strict control of the timing, all action-observation conditions were implemented using videos presented onscreen. Studies on the perception of robotic movement found that it can activate the mirror neuron system as well (Van Overwalle & Baetens, 2009), perhaps increasingly so with increased human-likeness (Matsuda et al., 2016) and familiarity (Press, 2011). To our knowledge, the only study using robotic stimuli in a study on (perceptual) SA in the auditory modality was Sato (2008). In this study, tones generated by a robotic arm were not attenuated compared to external tones, but tones that were self-generated or generated by an observed person were attenuated compared to the robot condition, thus suggesting that the agency attribution to a human agent is relevant (Sato, 2008).

Although the relation between perceptual and neurophysiological measures is unclear (e.g., Palmer et al., 2016), we nevertheless expected to find lower P2 amplitudes not only for tones that were self-generated, but also for tones that were generated by another human agent compared to those without a human agent, to visually cued tones and un-cued external tones (Egan et al., 2023; Ghio et al., 2021). Furthermore, we expected P2 attenuation for tones not generated by human agents, but which were yet temporally predictable, compared to the condition with unpredictable ones (Egan et al., 2023). Because some studies suggested that the N1 attenuation solely relates to the temporal predictability of the tone (e.g., Harrison et al., 2021), we expected N1 attenuation for all action-observation conditions and the cue condition in similar magnitude as for self-generation. Additionally, we acquired explicit measures of agency (i.e., judgement and confidence rating) to examine their influence on P2 amplitudes (Timm et al., 2016).

In additional analyses, we considered the modulation of the attenuation effects by certain individual traits of the participants measured via questionnaires. In previous studies, delusional ideation or anomalous perception was hypothesised to affect SA (Cao & Gross, 2015; Egan et al., 2023) and so we included the Cardiff anomalous perceptions scale (CAPS; Bell et al., 2006). Given its relation to the mirror neuron system (Iacoboni, 2009; but see Lamm & Majdandzic, 2015), empathy has been considered as a modulatory variable for SA in action observation (Cao & Gross, 2015; Egan et al., 2023), and so we also assessed the short form of the empathy quotient (EQ; Wakabayashi et al., 2006). The potential SA for tones generated by a robotic arm might further be modulated by the participants' tendency to anthropomorphise it. According to Epley et al. (2007), this tendency could be increased, for instance, by higher perceived similarity in motion and morphology, and possibly by individualism. Therefore, we included the anthropomorphism questionnaire (AQ; Neave et al., 2015). Participants with higher AQ scores might process the robotic action-observation conditions more similarly to the human action-observation condition than participants with lower scores. Lastly, we assessed the general attitudes towards robots scale (GAToRS; Koverola et al., 2022) to account for potential biases of participants. For instance, if participants have stronger negative attitudes, they might selectively anthropomorphise the robotic hand less than they would other objects (and show less attenuation).

Methods

Sample

To further investigate N1 and P2 attenuation for tones that were self-generated and generated by another participant based on the data from 36 (Ghio et al., 2018) and 38 participants (Ghio et al., 2021) and in particular the P2 attenuation effect that was found based on data from 49 participants in Egan et al. (2023) for tones that were self-generated, generated by another participant, or visually cued, we aimed to recruit a sample of similar size. In the current study, we acquired 49 healthy right-handed participants in total. However, nine participants were excluded from analysis because the data of one experimental block was missing (two due to technical difficulties, seven due to task-irrelevant button presses, i.e., not following instructions for that block). Therefore, the analysed sample consisted of 40 participants (10 male). The mean age was 23.63 years (*SD* = 4.88; range: 18 - 40). Before beginning the experiment participants gave written informed consent. They received up to 25 Euro or course credit for participation. The ethics committee of the Faculty of Mathematics and Natural Sciences at Heinrich Heine University Düsseldorf has approved this study.

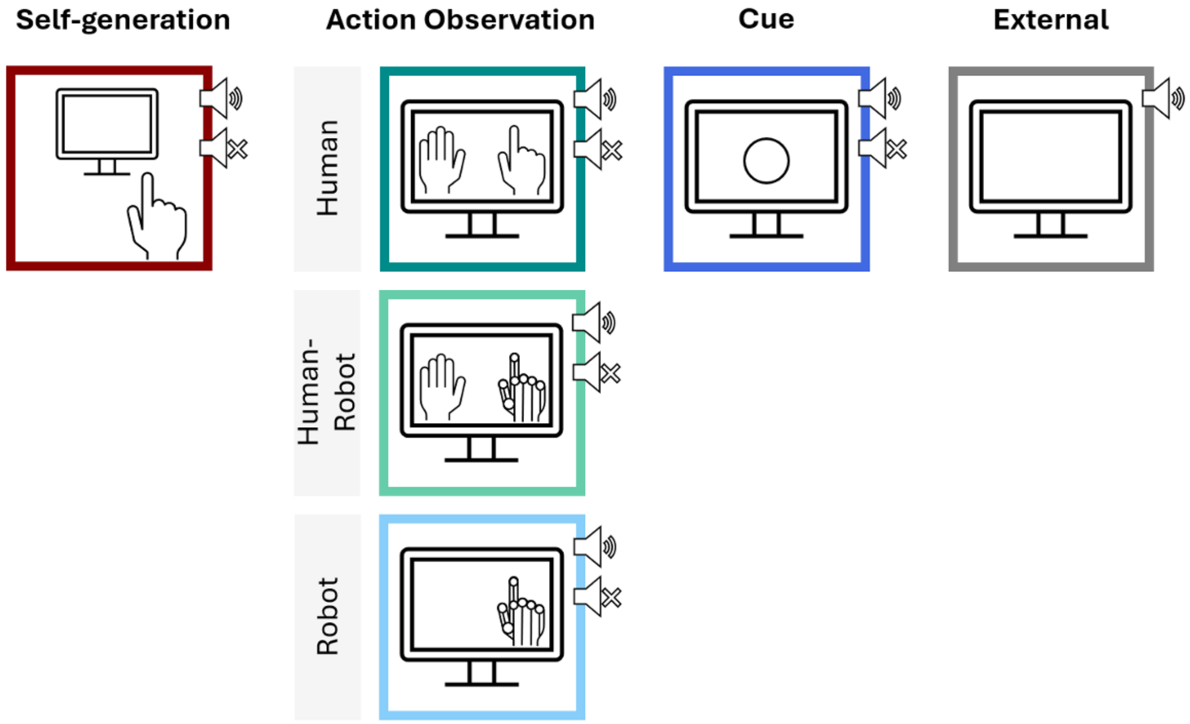
Experimental Design and Conditions

Based on the contingent paradigm (Horváth, 2015) and our previous studies that adapted the paradigm for action observation (Egan et al., 2023; Ghio et al., 2021; Ghio et al., 2018; Seidel et al., 2023), we included the self-generation and external-tone conditions as well as conditions with action observation and a condition in which a visual cue preceded the tone. Thus, we compared the processing of self-generated tones, tones that were temporally predictable due to an observed action, tones that were temporally predictable due to a visual cue, and external tones that were not temporally predictable. The latter represent the comparison condition to which SA was determined in previous studies (Horváth, 2015). Our study comprised three action-observation conditions: (1) Observation of button presses performed by a human agent (see Ghio et al., 2018; Seidel et al., 2023), and (2) observations of button presses performed by a robotic hand that was controlled by another person or (3) by a computer. The cue condition entailed the presentation of a white circle preceding each tone (see Egan et al., 2023). Thus, we explored whether the neurophysiological SA we previously found for action observation (Egan et al., 2023; Ghio et al., 2021; Ghio et al., 2018; Seidel et al., 2023) is modulated by the attribution of agency to another person, while accounting for possible effects of the action-effect association (i.e., button press and tone) and temporal predictability (cf. Egan et al., 2023, for live action observation).

For all conditions but the external-tone condition, there was an experimental block and an additional control block. In the experimental blocks, tones were time-locked to a preceding event, that is, a self-performed action, an observed action or visual cue. In the respective control block, the event took place without the presentation of tones. The purpose of these control blocks was to calculate the difference between the ERPs of events with and events without tones, and thus correct for activity elicited by the preceding event and thus unrelated to the tone (cf. motor correction in the contingent paradigm; Horváth, 2015). The experimental design is visualised in Figure 1.

Figure 1

Experimental Design



Note. Each condition is depicted and its related blocks are indicated with the sound-signs. For example, the self-generation condition consisted of one experimental block (in which the self-performed button press elicited a tone as indicated by the active sound-sign) and one control block (in which the self-performed button press elicited no tone as indicated by the muted sound-sign). The self-generation condition with its two blocks and the external condition were comparable to the contingent paradigm (Horváth, 2015). In the contingent paradigm, the block in which the self-performed button press elicits no tone serves to correct for button-press-related (and thus tone-unrelated) activity in the event-related potentials of the self-generated tones. We applied the same procedure for the action-observation and cue conditions, and therefore, included a control block for each.

Self-Generation and External Conditions

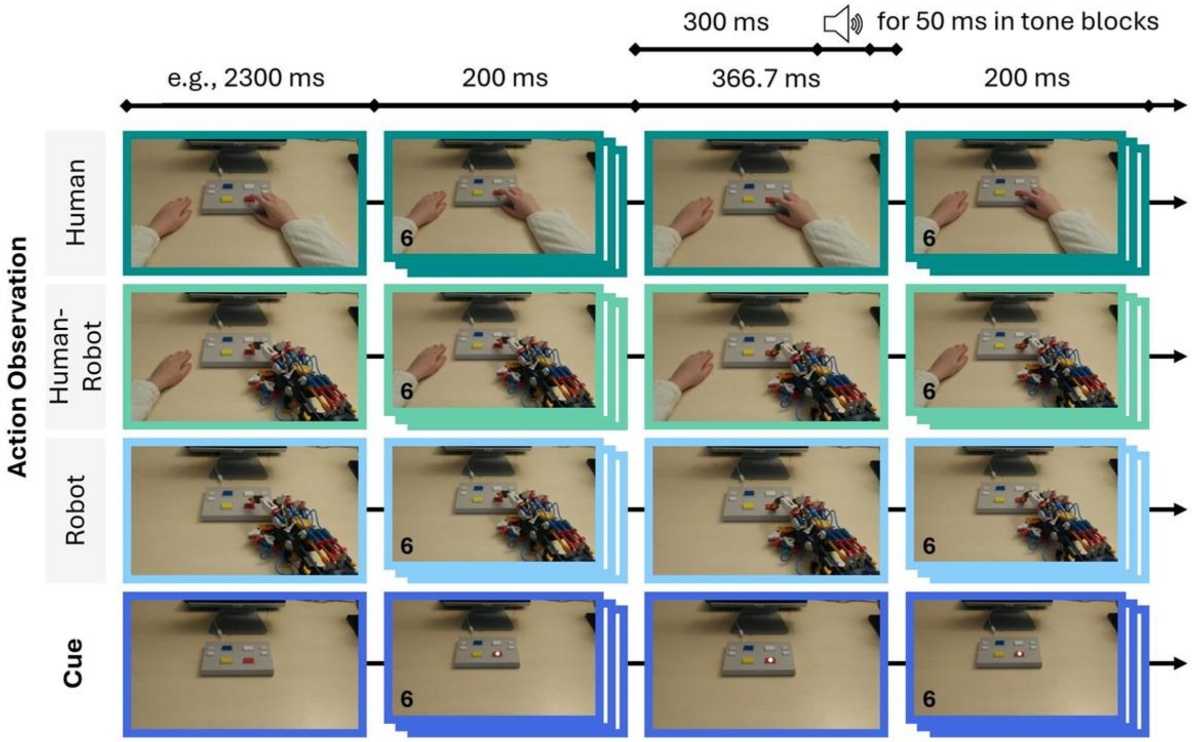
To determine SA, tone processing in the (corrected) self-generation condition (where the correction is the result of the difference of experimental and control blocks) is typically compared to tone processing in the external-tone condition (Horváth, 2015). In the experimental block of the self-generation condition, participants were instructed to press a particular button (the red button on the response box) in the learned rhythm of 2400 ms and the button press elicited a tone 300 ms after onset. In the control block, participants were also instructed to press the button in the learned rhythm, but no tone was elicited. In the external condition, the interval between the tones was determined by a function that picked a random number from a normal distribution with a mean of 2400 ms and a standard deviation of 350 ms (overall for the analysed sample: M = 2425 ms, SD = 42 ms). Participants were instructed to listen to the tones. In the self-generation experimental block, its control block, and the external condition, a picture of the response box was shown on the computer screen throughout, and participants were instructed to fixate the red button.

Action-Observation and Cue Conditions

There were three action-observation conditions (i.e., human, human-robot, and robot). In all three conditions, the button press (of the red button on the response box) was viewed from a first-person perspective and shown via a sequence of pictures (see Figure 2). In the human action-observation condition, participants observed a right human hand pressing the button. Additionally, a left human hand was visible on the side to further highlight that the agent was human. In the human-robot action-observation condition, participants observed a right robotic hand pressing the button, with a left human hand visible on the side, and were instructed that a person was controlling the hand (i.e., human agent with robotic depiction). In the robot-action observation condition, participants viewed a right robotic hand pressing the button without a depiction of a left hand and were instructed that the computer controlled it (i.e., computer as agent with robotic depiction). In the experimental blocks of all action-observation conditions, the tone was elicited 300 ms after the picture showing the button completely pressed (see Figure 2 for the timings) and instructions were to fixate the red button, observe the button press, and listen to the tones. Similar to the self-generation condition, there were also control blocks, in which observed button presses did not elicit tones. Similar to the external condition, the inter-trial-intervals for all action-observation conditions (experimental and control blocks) were determined by the function described above. However, due to a latency error of the stimulus-presentation monitor, each picture was presented for one more frame than intended and the intervals, which were calculated based on the a priori duration assumptions, were consistently longer for all blocks of the action-observation conditions (*M* = 2880-2896 ms, *SD* = 33-42 ms).

Figure 2

Trials in the Action-Observation and Cue Conditions



Note. The timings in the human, human-robot, and robot action-observation conditions were identical. After a random interval presenting the first picture, the sequential presentation of six pictures (each presented for 33 ms, i.e., two frames) led to the picture of the completely pressed button. During the presentation of this picture the tone was played. Then, a second picture sequence showing the same six pictures in reverse order led back to the first picture for the subsequent trial. In the cue condition, the timing was kept the same, though all six pictures in the videos and the picture shown while the tone was played were identical: a white cue on the response button.

In the cue condition, participants saw a white circle appear and disappear on the red button of the response box pictured onscreen that was pressed or observed to be pressed in other conditions (self-generation and action-observation). Timings were kept comparable to the action-observation conditions (see Figure 2). In the experimental block, the tone was presented 500 ms (i.e., the sum of the duration of the button-press motion and the action-outcome delay in the action-observation conditions) after the onset of the white circle. In the control block, the white circle was not followed by a tone. The inter-trial-intervals were again determined by the function described above and, as for the action-observation conditions, were consistently longer than in the external condition (*M* = 2880 and 2894 ms, *SD* = 42 and 43 ms).

Questionnaires

We included four questionnaires to assess inter-individual differences in traits between participants which could potentially modulate attenuation for self-performed action or action observation. The questionnaire on anomalous perception (CAPS; Bell et al., 2006) assessed a potential moderating factor for the magnitude of SA for self-generated tones. The CAPS has 32 items, based on which a total score (i.e., number of items the participants claimed to have experienced) and three rating scores were calculated. These ratings were done on 1-to-5 scales for distress, intrusiveness, and frequency; and the rating scores were calculated as the sum of these (items not endorsed by the participants had a value of 0). For example, one item was: “Do you ever notice that sounds are much louder than they normally would be?” If the participant endorsed the item, they rated how distressing this anomalous perceptual experience was, how distracting it was, and how often it occurred. For a potential modulation of the attenuation in action-observation conditions a questionnaire measure of trait empathy (22-item short form of the EQ; Wakabayashi et al., 2006) was considered. This yielded a single score, the EQ. Given the use of robotic depictions, we additionally included a measure of anthropomorphism (AQ; Neave et al., 2015), and a questionnaire that scores general attitudes towards robots (GAToRS; Koverola et al., 2022). For the AQ, a sum score for the two sub-scales regarding childhood-related items (10 items, e.g., "When I was a child, I made sure that when I put my toys away the ones who were friends were placed side by side") and items regarding current (i.e., adult) beliefs and behaviours (10 items; e.g., "I sometimes wonder if my computer deliberately runs more slowly after I have shouted at it"), and a total score were calculated (i.e., sum of the sub-scale scores). The GAToRS included four attitude factors and resulted in four scores with five items each: personal level positive (i.e., comfort and enjoyment around robots), personal level negative (i.e., unease and anxiety around robots), societal level positive (i.e., reasonable hopes about robots in general), and societal level negative (i.e., reasonable worries about robots in general).

To evaluate the stimuli of the action-observation conditions (see Appendix) participants were asked to rate them on four items of the anthropomorphism scale of the Godspeed questionnaire series (Bartneck et al., 2008). Three items pertained to the overall impression (fake vs. natural, machine- vs. humanlike, and artificial vs. lifelike). One item related to the observed movement (moving rigidly vs. moving elegantly). Each item was rated on a 5-point-Likert scale.

Experimental Procedure

The entire acquisition took about 2.5 h. Participants first filled out a demographic questionnaire and the questionnaires assessing the traits described above (see Questionnaires). Then, participants were prepared for EEG recording (see Data Recording) and were seated in the recording chamber in front of a 22-inch LCD computer screen (resolution: 1680*1050, refresh rate: 60 Hz) and a Cedrus RB-844 response box (Cedrus Corporation). Sennheiser (HD 201) headphones were placed on top of the

EEG cap before the experiment was started. The experiment was run with the Presentation software (Version 20.3, Build 02.25.19; Neurobehavioral systems).

Before running the experimental conditions, consisting of experimental and control blocks (see Experimental Design and Conditions), participants underwent training that was divided into two blocks. The first block of the training served to introduce participants to the rhythm that they later had to reproduce during the self-generation experimental block and the respective control block. Eighty tones were played every 2400 ms and participants were instructed to press the button simultaneously. The tone (680 Hz, 50 ms duration, increase of amplitude in the first and decrease in the last 5 ms) was identical throughout the experiment. In the second training block, participants were instructed to press the button in the learned rhythm to elicit a tone, and they received feedback if they were too fast (< 1800 ms) or too slow (> 3000 ms). The tone occurred 300 ms after button press onset. In this block, as in all subsequent blocks, three tones were presented in the 2400 ms intervals at the beginning to remind participants of the learned rhythm and each block contained 80 trials (i.e., 80 button presses in this block).

The following blocks belonged to the experimental conditions (see Experimental Design and Conditions) and could be categorised into experimental and control blocks. The order of experimental blocks (one per condition, six in total) was randomised, as was the order of control blocks (one per condition, except for the external condition, five in total, see Figure 1). However, participants were presented with either all experimental blocks first or all control blocks first. After each tone block, as an explicit measure of the sense of agency, participants responded to two questions. In an agency judgement, they were asked to judge who the agent of the tone was: themselves, another person, or the computer. This was included as an instruction check. Then, they rated how strongly they ascribed agency to the chosen agent on a 6-point-Likert scale (i.e., agency-confidence rating). At the end of the experiment, three trials of each of the three action-observation conditions were presented without tones and participants were asked to rate the observed hand on the four selected items of the Godspeed questionnaire series (Bartneck et al., 2008; see Questionnaires).

EEG Acquisition and Data Preparation

Data Recording

EEG data were recorded with BrainVision Recorder (Version 1.21.0402) and a BrainAmp MR Plus amplifier (Brain Products GmbH). Twenty-eight passive Ag/AgCl electrodes and an electrode at AFz serving as ground were placed on an Easycap electrode cap (see Figure 3B for electrode sites). Reference electrodes were placed on the mastoids. To record vertical and horizontal eye movements, further electrodes were placed at the cap positions F9, F10, and FP2, and under the right eye. For the recording, the sampling rate was set to 1000 Hz and notch and other software filters were disabled. Impedances were kept below 10 kΩ and direct-current corrections were applied online.

Pre-processing

With BrainVision Analyzer (BVA; Version 2.2.0.7383, Brain Products GmbH) the following pre-processing steps were done. We applied a direct current detrend and Butterworth zero phase filters with a low cut-off of 0.1 Hz, a high cut-off of 30 Hz, and a notch filter of 50 Hz. The BVA's ocular correction independent component analysis was applied (steps = 512, Infomax restricted biased) and components of blinking and eye movement artifacts were manually selected and removed. Because the EEG data of some participants clearly indicated that they had closed their eyes, the data was specifically screened to identify and mark intervals showing alpha activity on most electrodes and longer blinks or blink-like activity with the raw data inspection function in BVA. A second rater checked the validity of the markings. Segmentation was done in two steps to apply trial exclusion criteria on a wider pre-tone interval. First, segments beginning 400 ms before and ending 400 ms after tone markers were extracted. Note that tone markers were also sent in the control blocks, but the tones were muted. In this step, all segments with intervals marked in the raw data inspection were excluded as well as those, in which participants pressed the button, although, according to the instructions, they were not required to do so (i.e., in all blocks except the experimental block of the self-generation condition and its control block). Furthermore, only segments with preceding inter-event intervals between 1800 and 10000 ms were considered and, for the self-generation condition, segments of double presses (i.e., when a second button press followed the button press after less than 800 ms) were also excluded. In the second segmentation step, the segments were shortened, yielding segments beginning 100 ms before the tone markers with a total length of 500 ms. The first 100 ms were used for the subsequent baseline correction. For artifact rejection, we applied the automatic procedure in BVA: maximal allowed voltage step: 50 µV/ms, maximal allowed difference of values in 100-ms intervals: 100 µV, allowed amplitudes: -100 to 100 µV, lowest allowed activity in 100ms: 0.5 µV (see Table 1 for numbers of trials retained in each block). Averages across trials were calculated for control blocks for use in the correction procedure (see next section).

Table 1

Numbers of Trials after Exclusions

Condition	Block	*M*	*SD*	Minimum	Maximum
Self-generation	Experimental	75.60	4.79	63	79
	Control	76.20	3.76	73	79
Action-Observation Human	Experimental	76.53	4.66	60	79
	Control	77.45	1.72	73	79
Action-Observation Human-Robot	Experimental	76.95	3.52	66	79
	Control	77.62	2.33	67	79
Action-Observation Robot	Experimental	77.08	3.93	57	79
	Control	77.10	2.72	69	79
Cue	Experimental	75.72	5.07	62	79
	Control	77.38	2.48	69	79
External		72.90	7.01	37	79

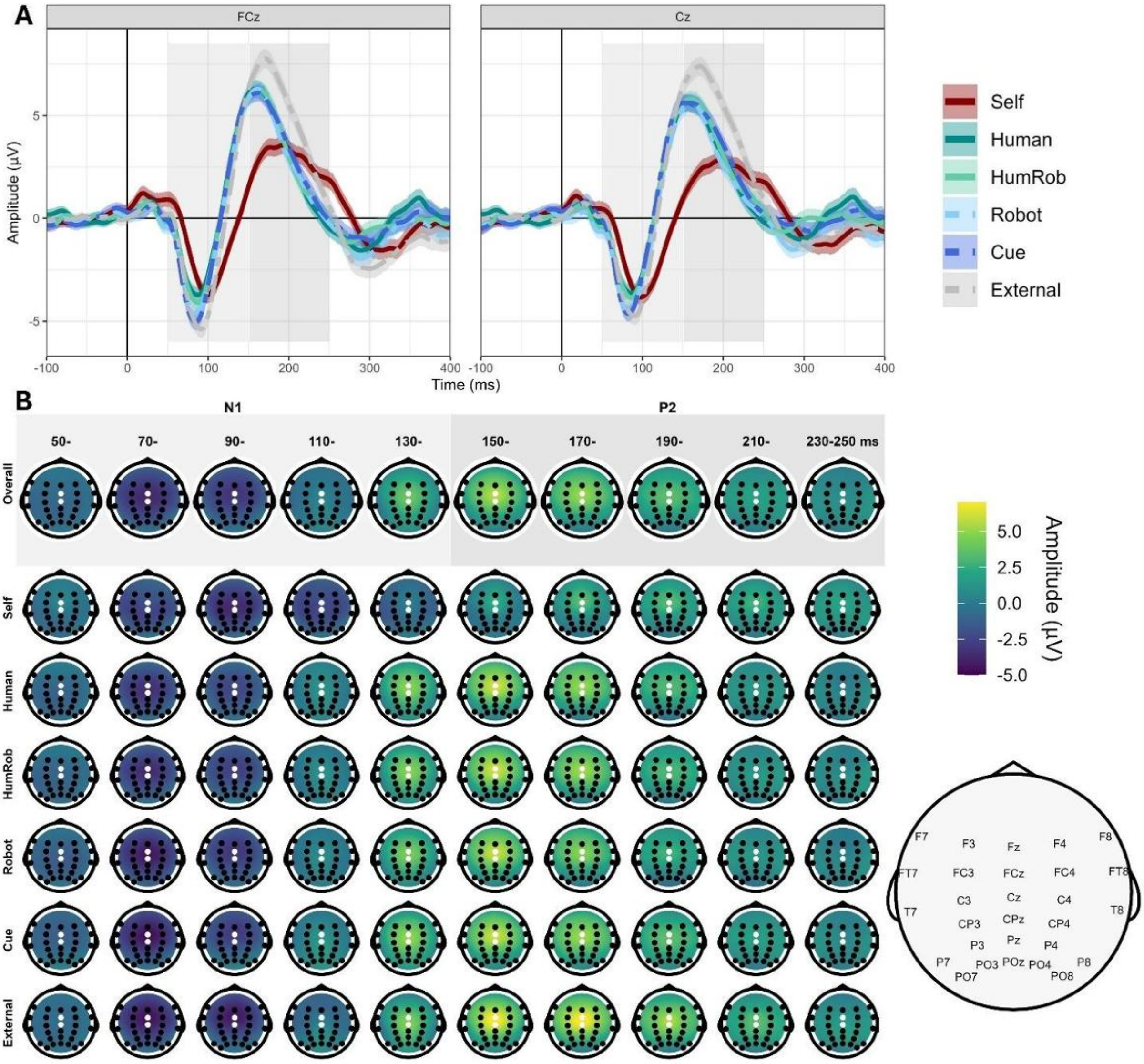
Note. All participants retained at least 70 % of trials (56 trials) in all blocks with one exception: 37 trials in the external condition due to a belated start of recording.

N1 and P2 Amplitude Extraction

For the remaining data preparation, we used R (Version 4.4.2, R Core Team, 2024). For the self-generation, action-observation, and cue conditions, the ERPs were corrected for tone-unrelated activity (cf. motor correction in contingent paradigm; Horváth, 2015). For this purpose, average ERPs of the control blocks were subtracted from the single-trial ERPs of the respective experimental blocks. Based on the topographies of the grand averages for the corrected ERPs of each condition, FCz and Cz were identified as electrodes with the most pronounced N1 and P2 peaks (see Figure 3B; cf. Ghio et al., 2018; Han et al., 2021; Harrison et al., 2021; Klaffehn et al., 2019) and chosen for further analysis. Because of latency differences between the conditions, N1 and P2 peaks were determined based on condition-specific grand averages of the corrected ERPs across participants (see Figure 3A). The N1 peak was determined as the local minimum in the interval between 50 and 150 ms for each condition (condition: FCz/Cz; self-generation: 99/98 ms, human: 87/86 ms, human-robot: 87/86 ms, robot: 83/82 ms, cue: 84/83 ms, external: 91/90 ms). Mean amplitudes were extracted from the 20-ms interval in each single trial, centred around the respective condition-specific peak latency (cf. Harrison et al., 2021; Jack et al., 2021; Neszmélyi et al., 2022; van Laarhoven et al., 2021). The P2 peak was determined as the local maximum in the interval between 150 and 250 ms for each condition (condition: FCz/Cz; self-generation: 192/195 ms, human: 158/155 ms, human-robot: 160/159 ms, robot: 162/161 ms, cue: 162/152 ms, external: 170/171 ms). As the P2 waveform was generally wider than the N1 (see Figures 3), mean amplitudes were extracted from the 40-ms interval in each single trial centred around the respective condition-specific peak latency (cf. Klaffehn et al., 2019; Pinheiro et al., 2019; van Laarhoven et al., 2021).

Figure 3

Condition-Specific Grand-Average ERPs and Topographies



Note. A: Grand-average ERPs for each condition and confidence intervals are plotted. This refers to the corrected ERPs for self-generation, action-observation (human, human-robot, and robot), and cue conditions. Grey areas mark the time intervals of the depiction of the topographies. B: Topographies of the grand average across all conditions and for the separate conditions are depicted in 20-ms bins over the interval between 50 and 250 ms, which contains N1 and P2 peaks. Electrodes FCz and Cz are highlighted in white. The plots were generated in R using the packages packages *ggplot2* (Wickham, 2016), *eegUtils* (Craddock, 2024) and *ggpubr* (Kassambara, 2023). Note that latency differences of N1 and P2 are visible between self, visual (action-observation [human, human-robot, robot] and cue), and external conditions. HumRob = Human-robot

Statistical Analysis

We used R for all statistical analyses, including the calculation of mixed-effects models. The effect of agent-condition (self-generation, human action-observation, human-robot action-observation, robot action-observation, cue, external with external as reference) was analysed for obtained agency-confidence ratings, and N1 and P2 amplitudes. Additionally, potential modulations of the N1 and P2 attenuation by various trait scores and by the agency-confidence ratings were tested.

Because of the single-item ratings of agency-confidence, we used the *ordinal* package (Christensen, 2023) to calculate a cumulative link mixed model on the ratings. For the N1 and P2 amplitudes, we used the *lme4* (Bates et al., 2015) and *lmerTest* packages (Kuznetsova et al., 2017) to calculate linear mixed models. These linear models were calculated with the optimiser *bobyca* and using the restricted maximum likelihood for model reporting and the maximum likelihood for model comparisons (Meteyard & Davies, 2020). For decomposition of fixed-effects contributions, the *anova* function was used for the mixed linear models (with a function to determine effect sizes by the package *effectsize*, Ben-Shachar et al., 2020). Estimated means and Bonferroni-corrected pairwise comparisons for significant effects were calculated with the package *emmeans* (Lenth, 2022). The package *performance* (Lüdecke et al., 2021) was used for model comparisons, which included likelihood ratio tests and Bayes Factors (BF).

Agency Judgements and Confidence Ratings

When exploring the data, we noted that, despite clear instructions, the agency judgements (i.e., identifying the agent as oneself, another person, or the computer) were not in all cases as intended (i.e., in the self-generation condition as oneself, in human and human-robot action-observation conditions as another person, and in robot action-observation, cue and external conditions as the computer). The number of participants who correctly judged the agent in all conditions was 17. On average, participants made 0.95 errors (*SD* = 0.99, range: 0-3). To determine whether these inaccuracies occurred consistently across conditions, we first computed chi-square tests for each condition with the observed frequencies of judgement responses (i.e., oneself, another person, and the computer) and adjusted expected frequencies (correct response: 0.9, incorrect responses each: 0.05). Significant results were obtained for action-observation conditions, that is, for the human, $\chi^2(2) = 38.50$, $p < .001$, human-robot, $\chi^2(2) = 57.01$, $p < .001$, and robot conditions, $\chi^2(2) = 22.47$, $p < .001$ (all other $p > .207$), indicating significantly fewer correct responses in these conditions. Then, a cumulative link mixed model for the judgement responses (i.e., self, other, and computer) was defined with only the intercept and random intercept by participant: Judgement ~ 1 + (1 | participant) and a model including also agent-condition as fixed effect. The model fit was improved by agent-condition, $\chi^2(5) = 260.51$, $p < .001$ (BF > 1000), indicating differences across conditions. Pairwise comparisons between conditions showed the expected significant differences resulting in condition clusters of external, cue and robot action-observation; human-robot and human action-observation; and self-generation. Thus, although the action-observation conditions (human, human-robot, and robot) showed increased inaccuracies in agency judgements, the overall pattern of judgements was as expected.

To analyse the 6-point-Likert scale agency-confidence ratings, a cumulative link mixed model was defined which, in a first step, comprised only the intercept and random intercept by participant: Rating ~ 1 + (1 | participant). The fixed effect agent-condition was added, Rating ~ agent-condition + (1 | participant), and the models were compared. The addition of agent-condition as a fixed effect improved the model fit, $\chi^2(5) = 27.62$, $p < .001$ (BF = 1.11); the addition of the judgement responses (i.e., oneself, another person, and the computer) as additional predictor of the ratings, $\chi^2(2) = 4.03$, $p = .133$ (BF = 0.03), or the addition of judgement accuracy (i.e., correct vs. incorrect), $\chi^2(1) = 0.03$, $p = .869$ (BF = 0.06), did not. Therefore, the model with only agent-condition as fixed effect was further analysed with pairwise comparisons.

Linear Mixed Models for N1 and P2 Amplitudes

To determine the random effect grouping factors, models were calculated with the maximum likelihood method and first defined with only the intercept as fixed effect and a random intercept by participant. For the N1, the addition of a random intercept by electrode to the model did not improve model fit, $\chi^2(1) = 1.56$, $p = .211$ (BF = 0.01). Adding agent-condition as fixed effect, $\chi^2(5) = 83.52$, $p < .001$ (BF > 1000), and adding a random by-participant slope for agent-condition, $\chi^2(20) = 738.77$, $p < .001$ (BF > 1000), greatly improved model fit. The final model, N1 ~ agent-condition + (agent-condition | participant), was recalculated with the restricted maximum likelihood method and further analysed with an analysis of variance and pairwise comparisons.

For the P2, the addition of a random intercept by electrode to the starting model improved model fit, $\chi^2(1) = 15.90$, $p < .001$ (BF = 14.84), as did the addition of agent-condition as fixed effect, $\chi^2(5) = 466.02$, $p < .001$ (BF > 1000), and the addition of a random by-participant slope for agent-condition, $\chi^2(20) = 870.88$, $p < .001$ (BF > 1000). The final model, P2 ~ agent-condition + (agent-condition | participant) + (1 | electrode), was recalculated with the restricted maximum likelihood method and further analysed with an analysis of variance and pairwise comparisons.

N1 and P2 Models Including Trait Scores

To investigate modulations of the N1 and P2 amplitude in the different agent-conditions by anomalous perception, empathy, anthropomorphism, and attitudes towards robots, model comparisons were planned. The N1 and P2 models resulting from the model building procedure described above (i.e., base model) were extended by one of the trait scores at a time (main and interaction effect). The base and extended models were then compared. For the N1, the trait models took the form: N1 ~ agent-condition*trait score + (agent-condition | participant), for the P2: P2 ~ agent-condition* trait score + (agent-condition | participant) + (1 | electrode). A total of 12 such extended models were calculated for each component. One model was calculated for empathy with the total score. For anomalous perception, the total score and three sub-scores (distress, intrusiveness, and frequency) were used in four separate models. Likewise, the total score and two sub-scores (childhood and current) were used for three models for anthropomorphism, and for the attitude towards robots, the four sub-scores (personal level positive, personal level negative, societal level positive, and societal level negative) were used in four separate models. For each trait score, the values were standardized by subtracting the mean and dividing by the standard deviation before adding the trait score to the basic model and comparing the models.

N1 and P2 Models Including Agency Ratings

To investigate modulations of the N1 and the P2 by agency-confidence ratings, further model comparisons were planned. Again, the N1 and P2 models from the model building procedure (see above) were extended and compared. However, due to the unexpected inaccuracies of agency judgements, not only the agency-confidence ratings and the interaction with agent-condition were added as fixed effects. A subsequent model included also judgement accuracy and all interactions. The models were compared sequentially. For the N1, the extended models were: N1 ~ agent-condition*agency-confidence rating + (agent-condition | participant) and N1 ~ agent-condition*agency-confidence rating*judgement accuracy + (agent-condition | participant). For the P2, the extended models were: P2 ~ agent-condition*agency-confidence rating + (agent-condition | participant) + (1 | electrode) and P2 ~ agent-condition*agency-confidence rating*judgement accuracy + (agent-condition | participant) + (1 | electrode).

Results

Agency-Confidence Ratings

Agency-confidence ratings (and agency judgments) are reported separately for each agent-condition in Table 2. Agency-confidence ratings were descriptively very high for self-generation, cue, and external conditions (each Mdn = 6; estimates in relation to external: 0.70 [*SE* = 0.49] for self, and 0.12 [*SE* = 0.48] for cue). Ratings were descriptively lower for the conditions for which more inaccurate agency judgements were found, namely, in the action-observation conditions: human (Mdn = 5; estimate in relation to external: -0.46, *SE* = 0.46), human-robot (Mdn = 4; estimate in relation to external: -1.43, *SE* = 0.46), and robot (Mdn = 5; estimate in relation to external: -0.85, *SE* = 0.46). Ratings of the human-robot action-observation condition were significantly lower than in the self-generation ($p < .001$), cue ($p = .011$), and external conditions ($p = .031$). The pairwise comparison between self-generation and robot action-observation conditions was also significant ($p = .014$). No other comparison reached significance (all *p*s > .20).

Table 2

Agency-Confidence Ratings

	Agent-Condition					
Rating	Self	Action-Observation			Cue	External
		Human	HumRob	Robot		
1	0 (0, 0, 0)	0 (0, 0, 0)	2 (0, 2, 0)	1 (0, 0, 1)	1 (0, 0, 1)	1 (0, 0, 1)
2	0 (0, 0, 0)	4 (0, 4, 0)	3 (0, 3, 0)	3 (0, 1, 2)	3 (0, 0, 3)	3 (0, 0, 3)
3	0 (0, 0, 0)	3 (0, 2, 1)	6 (0, 3, 3)	6 (0, 0, 6)	1 (0, 1, 0)	5 (0, 1, 4)
4	8 (7, 0, 1)	5 (0, 2, 3)	10 (0, 7, 3)	5 (1, 2, 2)	3 (0, 1, 2)	3 (0, 1, 2)
5	8 (8, 0, 0)	11 (0, 8, 3)	10 (0, 7, 3)	11 (0, 1, 10)	10 (0, 0, 10)	5 (0, 0, 5)
6	24 (24, 0, 0)	17 (0, 14, 3)	9 (0, 6, 3)	14 (0, 2, 12)	22 (1, 0, 21)	23 (2, 1, 20)

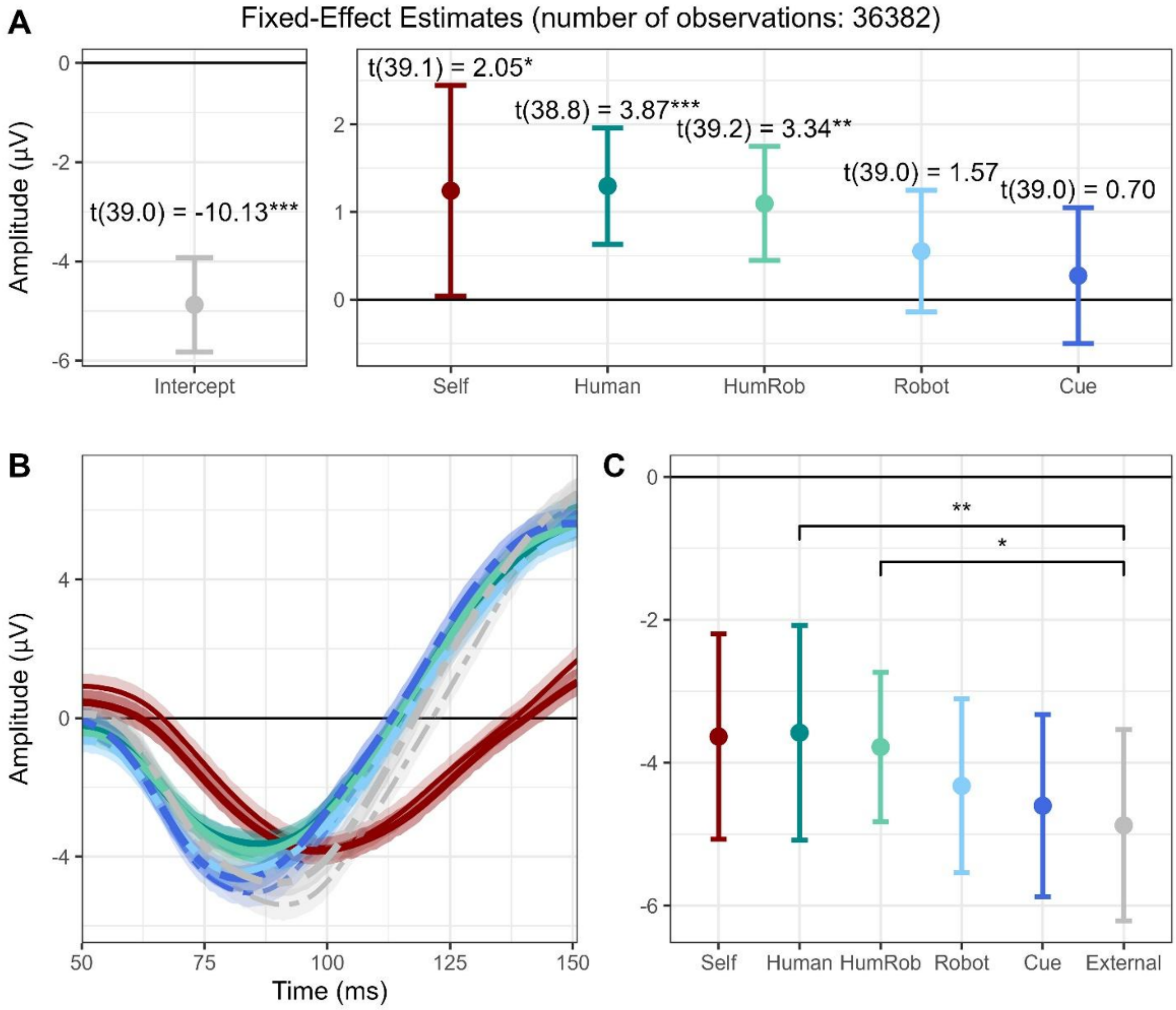
Note. The number of agency-confidence ratings for each level of the 6-point-Likert scale for each condition is listed. In brackets, the numbers for the respective agency judgement responses (oneself, another person, the computer) are reported. HumRob = Human-robot

N1 and P2 Models

The agent-condition ERPs of the entire segment are depicted in Figure 3A. For the N1, the effect of agent-condition was significant, $F(5, 38.95) = 6.20$, $p < .001$, $\eta^2_p = .44$. The fixed-effect estimates and the estimated means are depicted in Figure 4. Note that though the estimates for self-generation, human action-observation, and human-robot action-observation conditions are significantly different to the external condition (see Figure 4A), the Bonferroni-corrected pairwise comparisons to the external condition reached significance only for human (*p* = .006) and human-robot action-observation (*p* = .028), indicating significant attenuation of the N1 in these conditions (see Figure 4C). All other comparisons were not significant (all *ps* > .25).

Figure 4

N1 Results



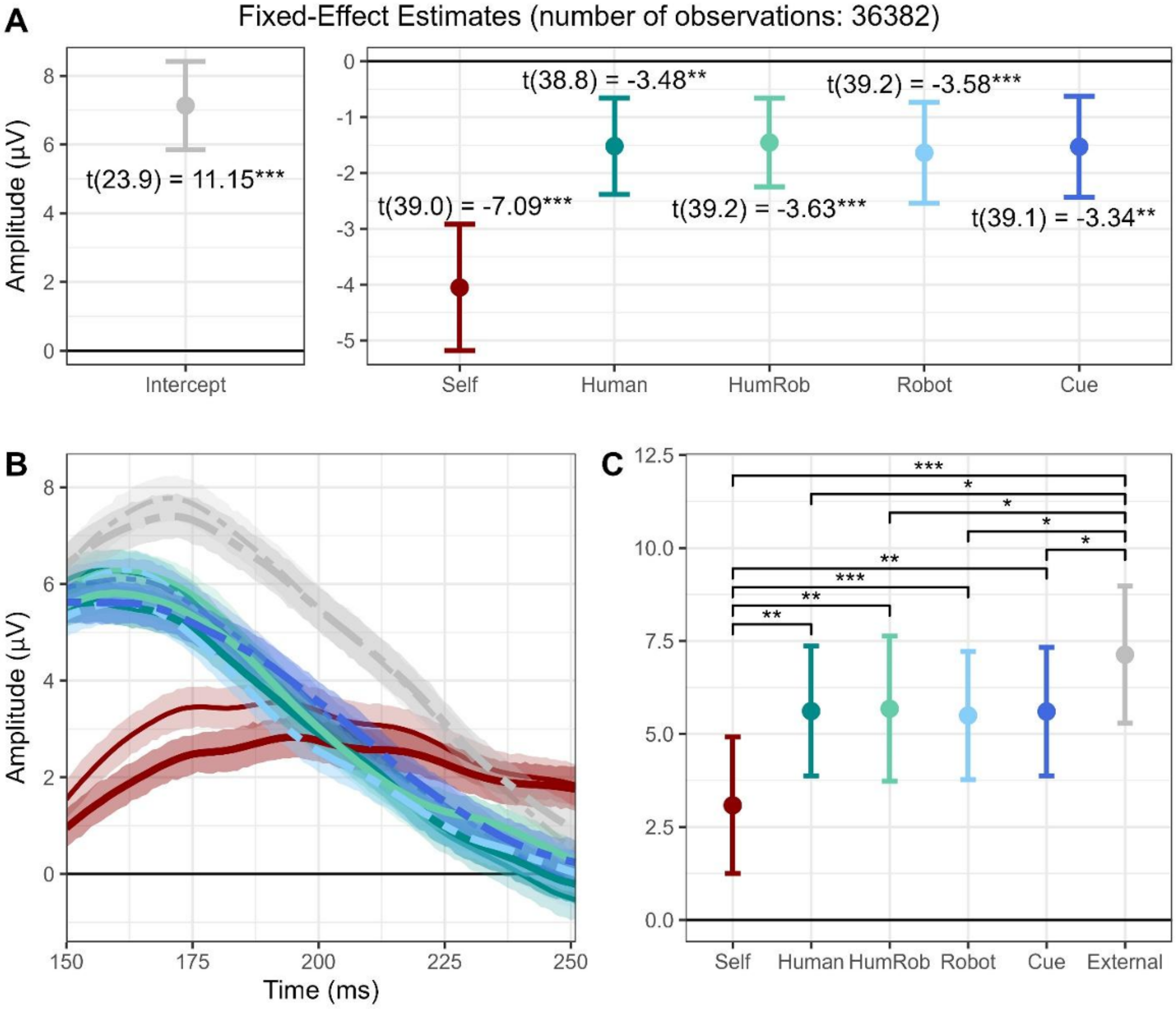
Note. The N1 model formula was: N1 ~ agent-condition + (agent-condition | participant). Agent-condition comprised self-generation (self), action-observation (human, human-robot [HumRob], and robot), cue and external conditions. A: Fixed-effect estimates of the model with confidence intervals.

The external condition was the reference. B: Condition-specific grand averages with confidence intervals. These are depicted for FCz (thin, light) and Cz (thick, dark). C: Estimated means with Bonferroni-adjusted confidence intervals. Additionally, significant Bonferroni-corrected pairwise comparisons are highlighted. The plots were generated in R using the packages *ggplot2* (Wickham, 2016), *ggpubr* (Kassambara, 2023), and *ggsignif* (Ahlmann-Eltze & Patil, 2021).
* < .050, ** < .010, *** < .001

For the P2, the effect of agent-condition was significant, $F(5, 38.77) = 14.00$, $p < .001$, $\eta^2_p = .64$. The fixed-effect estimates and the estimated means are depicted in Figure 5. Bonferroni-corrected pairwise comparisons showed significant attenuation to the external condition for the self-generation ($p < .001$), action-observation (human: $p = .019$, human-robot: $p = .012$, and robot: $p = .014$), and cue conditions ($p = .027$). Furthermore, the self-generation condition was significantly attenuated compared to all other conditions: human ($p = .006$), human-robot ($p = .005$), and robot action-observation ($p < .001$); and cue ($p = .005$). The other conditions did not differ significantly from each other ($p > .99$).

Figure 5

P2 Results



Note. The P2 model formula was: P2 ~ agent-condition + (agent-condition | participant) + (1 | electrode). Agent-condition comprised self-generation (self), action-observation (human, human-robot [HumRob], and robot), cue and external conditions. A: Fixed-effect estimates of the model with confidence intervals. The external condition was the reference. B: Condition-specific grand averages with confidence intervals. These are depicted for FCz (thin, light) and Cz (thick, dark). C: Estimated means with Bonferroni-adjusted confidence intervals. Additionally, significant Bonferroni-corrected pairwise comparisons are highlighted. The plots were generated in R using the packages *ggplot2* (Wickham, 2016), *ggpubr* (Kassambara, 2023), and *ggsignif* (Ahlmann-Eltze & Patil, 2021). HumRob = Human-robot

* < .050, ** < .010, *** < .001

Model Comparisons: Trait Scores

Comparisons of models for the N1 and the P2 including trait scores are reported in Table 3. For the N1, the addition of the GAToRS societal-negative score improved model fit according to the likelihood ratio test but not according to the BF. Similar results were found for the P2 amplitudes. The addition of the GAToRS societal-positive score improved model fit according to the likelihood ratio test but not according to the BF. For none of the other trait scores did the model fit improve for the N1 or P2 according to the likelihood ratio test or the BF. In sum, no model extended by an obtained trait score convincingly improved model fit for N1 or P2 amplitudes.

Table 3

Model Comparisons Including Trait Scores

Component	Model Extension	$\chi^2(6)$	*p*	BF
N1	Anomalous perception total	4.97	.547	< 0.01
	Distress	6.05	.418	< 0.01
	Intrusiveness	5.28	.508	< 0.01
	Frequency	7.23	.300	< 0.01
	Empathy	11.32	.079	< 0.01
	Anthropomorphism total	4.19	.650	< 0.01
	Childhood	3.19	.689	< 0.01
	Current	4.14	.657	< 0.01
	GAToRS			
	Personal-positive	5.95	428	< 0.01
	Personal-negative	4.88	.559	< 0.01
	Societal-positive	5.49	.482	< 0.01
	Societal-negative	24.40	< .001	< 0.01
P2	Anomalous perception total	7.59	.270	< 0.01
	Distress	7.99	.239	< 0.01
	Intrusiveness	6.98	.323	< 0.01
	Frequency	10.11	.120	< 0.01
	Empathy	11.59	.072	< 0.01
	Anthropomorphism total	6.20	.401	< 0.01
	Childhood	4.27	.640	< 0.01

Current	8.81	.185	< 0.01
GAToRS			
Personal-positive	4.67	.587	< 0.01
Personal-negative	2.83	.830	< 0.01
Societal-positive	13.43	.037	< 0.01
Societal-negative	8.06	.234	< 0.01

Note. For the N1 and P2, models including only the fixed effect of agent-condition were compared to the respective model extended by one of the 12 trait scores. Likelihood ratio tests (significant *p*-values indicating that the extended model is better) and Bayes Factors (BF; values greater than 1 indicating that the extended model is better) are reported. GAToRS = General attitudes towards robots scale

Model Comparisons: Agency Ratings

For the N1, the addition of the agency-confidence rating as predictor did not improve model fit, $\chi^2(6) = 4.84$, $p = .56$ (BF < 0.01), and the obtained model was not improved by the further addition of agency-judgement accuracy, $\chi^2(11) = 12.12$, $p = .36$ (BF < 0.01). For the P2, adding agency-confidence rating did not improve model fit either, $\chi^2(6) = 9.64$, $p = .14$ (BF < 0.01). Further adding agency judgement accuracy did improve model fit according to the likelihood ratio test, $\chi^2(11) = 23.58$, $p = .015$, but not according to the BF (BF < 0.01). In sum, the models including predictors related to explicit agency did not convincingly improve model fit for N1 or P2 amplitudes.

Discussion

In the present study, we explored effects of agency attribution on neurophysiological attenuation in action observation. This entailed comparisons of N1 and P2 amplitudes in response to tones that were generated by an observed action performed either by a human hand, a robotic hand controlled by a human agent, or a robotic hand controlled by a computer to respective ERP amplitudes for predictable tones cued by a circle or to unpredictable external tones. Additionally, tones generated by a self-performed action were included in the comparison. We expected to find P2 attenuation of temporally predictable tones compared to the unpredictable ones but stronger attenuation when the tones could be attributed to a human agent. As expected, P2 amplitudes in action-observation conditions and the cue condition were attenuated compared to the temporally unpredictable external tones. However, the P2 did neither differ between the action-observation conditions and the cue condition, nor among the different action-observation conditions. Furthermore, self-generated tones showed P2 attenuation not only compared to temporally unpredictable and cued (i.e., predictable) external tones, but also to the tones in the action-observation conditions. We further expected N1 attenuation of temporally predictable tones compared to the temporally unpredictable ones. Descriptively, self-generation and human action observation showed similar N1 amplitudes, and in human-robot action-observation, robot action-observation, cue, and external conditions N1 amplitudes were gradually more negative. However, significant attenuation relative to temporally unpredictable tones was found only for observation of a human and human-controlled robotic action (i.e., action-observation conditions with human agent). In additional analyses, we tested for modulations of N1 and P2 amplitudes by trait scores of anomalous perception, empathy, anthropomorphism, and attitude towards robots as well as by agency-confidence ratings. No such additional predictor significantly improved the models for N1 and P2 amplitudes.

Temporal Predictability and Action Observation

Temporal predictability is one of the confounds of the typical comparison of self-generated and un-cued external tones, as the former ones are temporally predictable and the latter ones are not (Hughes et al., 2013). Recent studies have found that temporal predictability might indeed explain part of the neurophysiological SA for the N1 (e.g., Harrison et al., 2021; but see Klaffehn et al., 2019) and the P2 component (e.g., Egan et al., 2023; but see Harrison et al., 2021). For the P2, the results of the present study support this explanation, in that the amplitudes for all temporally predictable tones were reduced relative to the unpredictable external tones, albeit with a retained self-generation effect (i.e., the attenuation for self-generated tones was strongest).

In contrast, the results for the N1 did not follow this pattern. Across human action-observation, human-robot action-observation, robot action-observation, and cueing, conditions had progressively lower values. Yet, the N1 amplitude was significantly reduced only for the first two conditions compared to the unpredictable external tones. Therefore, the results seem more in line with studies indicating no (e.g., Klaffehn et al., 2019) or a partial (e.g., Schafer & Marcus, 1973) effect of temporal predictability on the N1. In contrast, in a previous study, N1 amplitudes of tones generated by observed human action were only descriptively reduced and those of cued tones were significantly attenuated compared to un-cued external tones (Egan et al., 2023). However, in Egan et al. (2023) and the present study there was no significant difference in N1 amplitudes between action-observation and cueing conditions.

Whereas we expected this for the N1, we did not expect it for the P2. As mentioned in the introduction, previous studies found P2 attenuation in action observation similar to the one in self-generation for live action-observation settings (Egan et al., 2023; Ghio et al., 2021) but weaker for action observation via video (Ghio et al., 2018). Because we implemented action observation via video in the present study again, the weaker P2 attenuation for action observation compared to self-generation and absent attenuation relative to cued (i.e., predictable) tones thus might be a result of the action-observation setting. That is, when another person is observed onscreen the action-effect may show similar attenuation as any other temporally predictable stimulus. The reason for this modulation by observation setting could be a difference in the perception of agency, which might be diminished in the video setting.

P2 Attenuation and Agency

SA was previously related to the sense of agency (Hughes et al., 2013), particularly regarding the attenuation of the P2 component (e.g, Timm et al., 2016). Previous studies found P2 attenuation in action observation, possibly indicating a vicarious sense of agency (Egan et al., 2023; Ghio et al., 2021; Ghio et al., 2018). Our results replicated P2 attenuation in conditions with human agents other than oneself, but this was not significantly different to a condition containing a robotic hand without a human agent and to a condition containing only a visual circle as preceding event (i.e., controlling for temporal predictability). Importantly, the P2 amplitudes of self-generated tones were significantly attenuated compared to all these conditions, which was not the case in previous studies with live action observation (Egan et al., 2023; Ghio et al., 2021).

Self-agency might have a stronger influence on SA than vicarious agency. This was found for perceptual SA in two studies involving live action observation (Weiss et al., 2011a) and interactive button pressing with another person (Weiss et al., 2011b). This might be related to temporal control (i.e., controlling the onset time of the stimulus with an action), which is typically exclusive to the self-generation condition (Hughes et al., 2013). For neurophysiological SA, Harrison et al. (2021) found P2 attenuation only with temporal control (i.e., for self-timed button presses as opposed to button presses that were performed in response to a cue). However, effects of self-agency need not be limited to temporal control and explicit measures of agency can be considered. In this study, the agency judgements were indeed less accurate and agency-confidence ratings were lower in the action-observation conditions. However, the inclusion of explicit agency measures did not convincingly improve the P2 model. Thus, while the results of analyses of the explicit measures suggest weaker agency attribution could explain the lesser P2 attenuation in action-observation conditions, the measures did not sufficiently account for more variance, possibly due to the noise in data introduced by the inaccuracies and the constraints of the design: Clear instructions were given regarding the agent, perception of agency was not further manipulated (e.g., by varying action-outcome delays, Timm et al., 2016) and participants gave only one agency judgment and confidence rating per condition.

Future Perspectives

Another phenomenon regarded as an implicit measure of the sense of agency is intentional binding, which describes the shift of perceived timings of action and effect towards each other (for a review see Moore & Obhi, 2012). Vicarious sense of agency was found also in paradigms of intentional binding, in that the effect was found for action observation (e.g., Poonian & Cunnington, 2013). Studies including robotic agents found that in physical settings, as opposed to virtual settings, human-likeness was a relevant factor for finding the effect as well as action-occurrence (i.e., the robot performing a physical action, rather than a mere command transmission highlighted with visual cues on the robot; Roselli et al., 2022; Roselli et al., 2021). Further, a significant correlation of the perceived timing shift with intentionality attribution to the robot was found (Roselli et al., 2022; Roselli et al., 2021).

Although intentional binding likely results from different mechanisms than SA, as it might be less related to predictability and more to temporal control (Hughes et al., 2013), these findings offer avenues for alternative explanations and further research. Firstly, the naturalness, human-likeness, and lifelikeness were rated for our action-observation stimuli and the scores were found to be significantly lower for the robotic hand than for the human hand (see Appendix). Yet, the P2 amplitudes in these conditions did not differ significantly. Considering the P2 results in previous studies with live observation (Egan et al., 2023; Ghio et al., 2021) and the intentional binding results (Roselli et al., 2022), the P2 might only reflect such ratings, if the action observation were realised in a live setting instead of onscreen, in particular when using a robotic trigger that does not resemble a human hand. For instance, one study implemented a robotic agent as a sphere in virtual reality (Feder et al., 2023). However, the N1 of self-generated tones was attenuated compared to the robot-generated ones and the P2 was not, which is likely due to the lack of temporal predictability in the robot condition and the use of various action-outcome delays (Feder et al., 2023). Secondly, the smoothness of movement of our action-observation stimuli was rated and, although robotic-hand movement was rated to be significantly more rigid than the human-hand movement, the ratings for all stimuli were relatively low (i.e., rated as moving rigidly). This was due to the latency error of the stimulus-presentation monitor and might have led to lower perceived similarity in motion (less anthropomorphism) and less mirror neuron activation for human and robotic stimuli (Epley et al., 2007). Given this limitation and the onscreen presentation (instead of live observation), intentionality attribution and attribution of agency might have been impaired for all action-observation conditions in our study, yielding the significant P2 difference to self-generation and lack of significant P2 difference to cueing. To test this in future studies, the timings could be adjusted to increase similarity in motion to self-performed actions and a version of the experiment could be implemented in a live observation setting.

Limitations

Notably, the neurophysiological SA (i.e., attenuation of self-generated tones compared to un-cued external tones) that is commonly found (Horváth, 2015), was replicated in the present study only for the P2 component but not for the N1. Likewise, the N1 of self-generated tones in our previous study was only descriptively but not significantly attenuated (Egan et al., 2023). We attributed the absence of N1 attenuation in that study to a methodological deviation from the contingent paradigm (i.e., using a mirror construction to reflect the first-person perspective on the button press at an elevated position). However, this was not applied in the present study. Looking at the uncorrected contrasts for the fixed effects estimates in the present study, there was significant N1 attenuation for self-generation, but the confidence interval was comparatively large. The increased variance in N1 amplitude of self-generated tone might be due to additional attentional demands of the button-pressing task. Alternatively, a study introducing coincidental tones while participants performed button presses found an attenuating effect of temporal proximity on N1 and P2 amplitudes (Horváth et al., 2012). The lack of N1 attenuation might therefore also relate to the implemented action-effect delay of 300 ms. However, previous neurophysiological studies made use of longer action-effect delays and yet reported significant N1 attenuation (e.g., Klaffehn et al., 2019; Lange, 2011). Importantly, the N1 is modulated by inter-trial-intervals (SanMiguel et al., 2013), which in this study were longer in the self-generation condition (i.e., around 200 ms) and in the action-observation and cueing conditions (i.e., around 400 ms) than in the external condition. The standard deviation of intervals aggregated across participants were also larger in the self-generation condition (i.e., around 400 ms) than in all other conditions (i.e., 40 ms). It is possible that the amplitude increases by the higher inter-trial-intervals masked the attenuation that the self-generated and the other temporally predictable external tones might have had.

As mentioned above, there was a difference in attentional demands between conditions. The inaccurate agency judgements, the exclusions of participants because they pressed buttons when they were not instructed to do so, and the exclusions of trials because participants closed their eyes were all indications of a lack of attention, possibly due to the length of the experiment. While the order of blocks was randomised, effects of tiredness may especially affect blocks that did not require actions from the participants, such as the action-observation conditions.

As in our previous study (Egan et al., 2023), we aimed to examine the influence of inter-individual differences assessed by questionnaires on neurophysiological SA. Extending N1 and P2 models by trait scores, we again did not find any significant improvement of the models and thus no convincing indication of attenuation modulations by traits of participants (Egan et al., 2023). Measuring perceptual SA, Cao and Gross (2015) likewise found no correlation with empathy scores, but, in contrast to our study, did find a correlation of SA values to trait scores measuring delusional ideation. It is possible that the sample of the present study was too small to detect such modulations by trait in healthy participants, that the applied questionnaires were not suitable, or that other experimental manipulations are needed to investigate the potential modulations of SA by empathy, anomalous perception, and anthropomorphism.

Conclusion

In the present study, we explored the role of agency attribution to a human agent for neurophysiological SA during action observation, while accounting for effects of action-effect association and temporal predictability. Our results suggest that attenuation of N1 amplitudes is a less robust finding than that of P2 amplitudes and might be influenced by temporal predictability but cannot be fully explained by it. For the P2, there seemed to be an effect of temporal predictability and an additive effect of self-generation. For both components no effects of agency attribution were found. However, we discussed how the observation setting might affect attenuation in action observation, which could be addressed in future studies. The present study supports the notion of a combination of predictive mechanisms yielding neurophysiological SA (Dogge et al., 2019; Horváth, 2015; Hughes et al., 2013).

Acknowledgements

We would like to thank Hannah Kirchberg for her help in data acquisition.

Data Availability

Data and code are available at OSF repository:

https://osf.io/r4k63/?view_only=9a7cb312250d4c358db46872f4c856fe

References

Ahlmann-Eltze, C., & Patil, I. (2021). *ggsignif: R Package for Displaying Significance Brackets for 'ggplot2'*. In https://psyarxiv.com/7awm6

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2008). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, *1*(1), 71-81. https://doi.org/10.1007/s12369-008-0001-3

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). The Cardiff Anomalous Perceptions Scale (CAPS): A New Validated Measure of Anomalous Perceptual Experience. *Schizophrenia Bulletin*, *32*(2), 366-377. https://doi.org/10.1093/schbul/sbj014

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standatdized Parameters. *Journal of Open Source Software*, *5*(56), 2815. https://doi.org/10.21105/joss.02815

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *NeuroReport*, *12*(9), 1879-1884. https://doi.org/10.1097/00001756-200107030-00023

Cao, L., & Gross, J. (2015). Cultural Differences in Perceiving Sounds Generated by Others: Self Matters. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01865

Christensen, R. H. B. (2023). *ordinal-Regression Models for Ordinal Data*. In https://CRAN.R-project.org/package=ordinal

Craddock, M. (2024). *eegUtils: Utilities for Electroencephalographic (EEG) Analysis*. In https://github.com/craddm/eegUtils

Desantis, A., Weiss, C., Schutz-Bosbach, S., & Waszak, F. (2012). Believing and perceiving: authorship belief modulates sensory attenuation. *PLOS ONE*, *7*(5), e37959. https://doi.org/10.1371/journal.pone.0037959

Dogge, M., Custers, R., & Aarts, H. (2019). Moving forward: On the Limits of Motor-Based Forward Models. *Trends in Cognitive Sciences*, *23*(9), 743-753. https://doi.org/10.1016/j.tics.2019.06.008

Egan, S., Ghio, M., & Bellebaum, C. (2023). Auditory N1 and P2 attenuation in action observation: An event-related potential study considering effects of temporal predictability and individualism. *Biological Psychology*, *180*, 108575. https://doi.org/10.1016/j.biopsycho.2023.108575

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864-886. https://doi.org/10.1037/0033-295X.114.4.864

Feder, S., Miksch, J., Grimm, S., Krems, J. F., & Bendixen, A. (2023). Using event-related brain potentials to evaluate motor-auditory latencies in virtual reality. *Frontiers in Neuroergonomics*, *4*, 1196507. https://doi.org/10.3389/fnrgo.2023.1196507

Ferrari, P. F., & Rizzolatti, G. (2014). Mirror neuron research: the past and the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130169. https://doi.org/10.1098/rstb.2013.0169

Ford, J. M., & Mathalon, D. H. (2012). Anticipating the future: automatic prediction failures in schizophrenia. *International Journal of Psychophysiology*, *83*(2), 232-239. https://doi.org/10.1016/j.ijpsycho.2011.09.004

Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Daléry, J., d'Amato, T., & Jeannerod, M. (2001). Defective Recognition of One's Own Actions in Patients With Schizophrenia. *American Journal of Psychiatry*, *158*(3), 454-459. https://doi.org/10.1176/appi.ajp.158.3.454

Ghio, M., Egan, S., & Bellebaum, C. (2021). Similarities and differences between performers and observers in processing auditory action consequences: Evidence from simultaneous EEG acquisition. *Journal of Cognitive Neuroscience*, *33*(4), 683-694. https://doi.org/10.1162/jocn_a_01671

Ghio, M., Scharmach, K., & Bellebaum, C. (2018). ERP correlates of processing the auditory consequences of own versus observed actions. *Psychophysiology*, *55*(6), e13048. https://doi.org/10.1111/psyp.13048

Han, N., Jack, B. N., Hughes, G., Elijah, R. B., & Whitford, T. J. (2021). Sensory attenuation in the absence of movement: Differentiating motor action from sense of agency. *Cortex*, *141*, 436-448. https://doi.org/10.1016/j.cortex.2021.04.010

Harrison, A. W., Mannion, D. J., Jack, B. N., Griffiths, O., Hughes, G., & Whitford, T. J. (2021). Sensory attenuation is modulated by the contrasting effects of predictability and control. *NeuroImage*, *237*, 118103. https://doi.org/10.1016/j.neuroimage.2021.118103

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of Imprecise Corollary Discharge in Schizophrenia to Auditory Hallucinations. *Archives of General Psychiatry*, *64*(3), 286-296. https://doi.org/10.1001/archpsyc.64.3.286

Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54-65. https://doi.org/10.1016/j.brainres.2015.03.038

Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action–sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience*, *24*(9), 1919-1931. https://doi.org/10.1162/jocn_a_00215

Hughes, G., Desantis, A., & Waszak, F. (2013). 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. https://doi.org/10.1037/a0028566

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, *60*, 653-670. https://doi.org/10.1146/annurev.psych.60.110707.163604

Jack, B. N., Chilver, M. R., Vickery, R. M., Birznieks, I., Krstanoska-Blazeska, K., Whitford, T. J., & Griffiths, O. (2021). Movement Planning Determines Sensory Suppression: An Event-related Potential Study. *Journal of Cognitive Neuroscience*, *33*(12), 2427-2439. https://doi.org/10.1162/jocn_a_01747

Kassambara, A. (2023). *ggpubr: 'ggplot2' Based Publication Ready Plots*. In https://CRAN.R-project.org/package=ggpubr

Kiepe, F., Kraus, N., & Hesselmann, G. (2021). Sensory Attenuation in the Auditory Modality as a Window Into Predictive Processing. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.704668

Klaffehn, A. L., Baess, P., Kunde, W., & Pfister, R. (2019). Sensory attenuation prevails when controlling for temporal predictability of self-and externally generated tones. *Neuropsychologia*, *132*, 107145. https://doi.org/10.1016/j.neuropsychologia.2019.107145

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*(9), 2449-2461. https://doi.org/10.1016/j.cortex.2012.12.012

Korka, B., Widmann, A., Waszak, F., Darriba, Á., & Schröger, E. (2022). The auditory brain in action: Intention determines predictive processing in the auditory system—A review of current paradigms and findings. *Psychonomic Bulletin & Review*, *29*(2), 321-342. https://doi.org/10.3758/s13423-021-01992-z

Koverola, M., Kunnari, A., Sundvall, J., & Laakasuo, M. (2022). General Attitudes Towards Robots Scale (GAToRS): A New Instrument for Social Surveys. *International Journal of Social Robotics*, *14*(7), 1559-1581. https://doi.org/10.1007/s12369-022-00880-3

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Lamm, C., & Majdandzic, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy - A critical comment. *Neuroscience Research*, *90*, 15-24. https://doi.org/10.1016/j.neures.2014.10.008

Lange, K. (2011). The reduced N1 to self-generated tones: an effect of temporal predictability? *Psychophysiology*, *48*(8), 1088-1095. https://doi.org/10.1111/j.1469-8986.2010.01174.x

Lenth, R. V. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. In https://CRAN.R-project.org/package=emmeans

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Matsuda, G., Hiraki, K., & Ishiguro, H. (2016). EEG-Based Mu Rhythm Suppression to Measure the Effects of Appearance and Motion on Perceived Human Likeness of a Robot. *Journal of Human-Robot Interaction*, *5*(1), 68-81. https://doi.org/10.5898/JHRI.5.1.Matsuda

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*. https://doi.org/10.1016/j.jml.2020.104092

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and Cognition*, *21*(1), 546-561. https://doi.org/10.1016/j.concog.2011.12.002

Neave, N., Jackson, R., Saxton, T., & Hönekopp, J. (2015). The influence of anthropomorphic tendencies on human hoarding behaviours. *Personality and Individual Differences*, *72*, 214-219. https://doi.org/10.1016/j.paid.2014.08.041

Neszmélyi, B., Nevelo, D., & Horváth, J. (2022). Turning a blind eye to motor differences leads to bias in estimating action-related auditory ERP attenuation. *Biological Psychology*, *173*, 108387. https://doi.org/10.1016/j.biopsycho.2022.108387

Palmer, C. E., Davare, M., & Kilner, J. M. (2016). Physiological and Perceptual Sensory Attenuation Have Different Underlying Neurophysiological Correlates. *Journal of Neuroscience*, *36*(42), 10803-10812. https://doi.org/10.1523/JNEUROSCI.1694-16.2016

Pinheiro, A. P., Schwartze, M., Gutierrez, F., & Kotz, S. A. (2019). When temporal prediction errs: ERP responses to delayed action-feedback onset. *Neuropsychologia*, *134*, 107200. https://doi.org/10.1016/j.neuropsychologia.2019.107200

Poonian, S. K., & Cunnington, R. (2013). Intentional binding in self-made and observed actions. *Experimental Brain Research*, *229*(3), 419-427. https://doi.org/10.1007/s00221-013-3505-5

Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. *Neuroscience and Biobehavioral Reviews*, *35*(6), 1410-1418. https://doi.org/10.1016/j.neubiorev.2011.03.004

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. In R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Roselli, C., Ciardo, F., De Tommaso, D., & Wykowska, A. (2022). Human-likeness and attribution of intentionality predict vicarious sense of agency over humanoid robot actions. *Scientific Reports*, *12*(1), 13845. https://doi.org/10.1038/s41598-022-18151-6

Roselli, C., Ciardo, F., & Wykowska, A. (2021). Intentions with actions: The role of intentionality attribution on the vicarious sense of agency in Human-Robot interaction. *Quarterly Journal of Experimental Psychology*, *75*(4), 616-632. https://doi.org/10.1177/17470218211042003

SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N 1 component of the auditory ERP. *Psychophysiology*, *50*(4), 334-343. https://doi.org/10.1111/psyp.12024

Sato, A. (2008). Action observation modulates auditory perception of the consequence of others' actions. *Consciousness and Cognition*, *17*(4), 1219-1227. https://doi.org/10.1016/j.concog.2008.01.003

Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain Responses. *Science*, *181*(4095), 175-177. https://doi.org/10.1126/science.181.4095.175

Seidel, A., Ghio, M., Studer, B., & Bellebaum, C. (2021). Illusion of control affects ERP amplitude reductions for auditory outcomes of self-generated actions. *Psychophysiology*, *58*(5), e13792. https://doi.org/10.1111/psyp.13792

Seidel, A., Weber, C., Ghio, M., & Bellebaum, C. (2023). My view on your actions: Dynamic changes in viewpoint-dependent auditory ERP attenuation during action observation. *Cognitive, Affective, & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-023-01083-7

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*(7), 1481-1489. https://doi.org/10.1162/jocn_a_00552

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. https://doi.org/10.1016/j.cortex.2016.03.018

van Laarhoven, T., Stekelenburg, J. J., & Vroomen, J. (2021). Suppression of the auditory N1 by visual anticipatory motion is modulated by temporal and identity predictability. *Psychophysiology*, *58*(3). https://doi.org/10.1111/psyp.13749

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, *48*(3), 564-584. https://doi.org/10.1016/j.neuroimage.2009.06.009

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., & Weil, L. (2006). Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). *Personality and Individual Differences*, *41*(5), 929-940. https://doi.org/10.1016/j.paid.2006.03.017

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011a). The self in action effects: Selective attenuation of self-generated sounds. *Cognition*, *121*(2), 207-218. https://doi.org/10.1016/j.cognition.2011.06.011

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011b). The Self in Social Interactions: Sensory Attenuation of Auditory Action Effects Is Stronger in Interactions with Others. *PLOS ONE*, *6*(7), 1-3. https://doi.org/10.1371/journal.pone.0022723

Weller, L., Schwarz, K. A., Kunde, W., & Pfister, R. (2017). Was it me? – Filling the interval between action and effects increases agency but not sensory attenuation. *Biological Psychology*, *123*, 241-249. https://doi.org/10.1016/j.biopsycho.2016.12.015

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. In Springer-Verlag New York. https://ggplot2.tidyverse.org

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 593-602. https://doi.org/10.1098/rstb.2002.1238

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729-R732. https://doi.org/10.1016/S0960-9822(01)00432-8

Appendix

Creation and Evaluation of Visual Stimuli

Creation of Visual Stimuli

Setup and Recording

The robotic hand used for the stimuli came from a KOSMOS experiment box for children with a hydraulic system. High-resolution videos of button presses, performed by a human hand and using the robotic hand, were recorded from a perspective similar to that of participants in the study. For the videos of the button press by the human hand, the other hand was placed on the table as well. For the videos of the button press by the robotic hand, it was not.

Frame Selection and Modification

For the human button press and the robotic button press eight frames were exported from the videos. The choice was based on the fluidity of the motion in the resulting videos and on comparability of the human and robotic motion. Because the robotic hand was larger than the human hand, only the relative motion was comparable. The frames of the human button press were used in the human action-observation condition. A frame from the video without hands was used as stationary visual background stimulus in self-generation and external condition, and with the addition of a white circle in the cue condition. Using the software GIMP (Version 2.10.32) the frames of the robotic button press were altered, so that the right human hand seen pulling the levers to control the robotic hand was removed. These frames were used in the robot action-observation condition. Then, the resting left hand from the human button press frames was inserted in each frame of the robotic button press. These frames were used in the human-robot action-observation condition.

Saving as Video

The six frames showing the pressing of the button were added together to a video using the Windows function available for this purpose and converted to an AVI file. This was done twice for each button press: in proper order showing the pressing of the button, and in reverse order showing the releasing of the button. It was also done in the exact same way with the cue picture (i.e., white circle on response button; i.e., using the same picture six times in the video) to ensure that the timing of the visual stimuli in observation and cue conditions would be identical.

Anthropomorphism Evaluation of Action-Observation Stimuli

Statistical Analyses

The evaluation was done based on selected items of the anthropomorphism scale from the Godspeed questionnaire series (Bartneck et al., 2008) for the visual stimuli used in the human, human-robot, and robot action-observation conditions. The purpose was to validate higher ratings in human compared to the robot action-observation conditions, and to get an estimate of anthropomorphism of the robotic hand we used for stimuli creation. Thus, the agent-condition factor in this analysis was reduced to the three action-observation conditions. For each of the obtained Godspeed anthropomorphism ratings (fake - natural, machinelike - humanlike, artificial - lifelike, moving rigidly - moving elegantly), a model was defined as specified for the main analyses with agent-condition as fixed effect and random intercept by participant. Bonferroni-corrected pairwise comparisons were calculated to detect significant differences between the conditions.

Results

All Godspeed anthropomorphism ratings showed the same pattern: the human action-observation condition had higher ratings than the human-robot and robot action-observation conditions (see Table A1). For the natural ratings, the human condition had a median of 4, and the human-robot condition a significantly lower score (Mdn = 1; estimate in relation to human: -2.80, *SE* = 0.53; pairwise comparison: *p* < .001), as did the robot condition (Mdn = 1; estimate in relation to human: -3.95, *SE* = 0.63; pairwise comparison: *p* < .001). The human-robot and robot conditions did not differ significantly (*p* = .065). Similarly, the human condition had a median humanlike rating of 4, which was significantly higher than the ratings in human-robot (Mdn = 1; estimate in relation to human: -3.56, *SE* = 0.66; pairwise comparison: *p* < .001) and robot conditions (Mdn = 1; estimate in relation to human: -4.91, *SE* = 0.83; pairwise comparison: *p* < .001). The latter two did not differ significantly (*p* = .088). For lifelike ratings, the human condition had a median of 3.5, which was significantly higher than the ratings in human-robot (Mdn = 1; estimate in relation to human: -3.61, *SE* = 0.66; pairwise comparison: *p* < .001) and robot conditions (Mdn = 1; estimate in relation to human: -5.95, *SE* = 0.85; pairwise comparison: *p* < .001). The latter two did not differ significantly (*p* = .051). Lastly, the median movement rating was 2 for each condition but ratings were significantly higher in the human condition than in human-robot (estimate in relation to human: -1.15, *SE* = 0.45; pairwise comparison: *p* = .035) and robot conditions (estimate in relation to human: -1.43, *SE* = 0.47; pairwise comparison: *p* = .008). Again, the human-robot and robot conditions did not differ significantly in their ratings (p > .999).

Table A1

Godspeed Ratings

Rating	Action-Observation Condition	1	2	3	4	5
Natural	Human	3	6	8	7	16
	Human-Robot	21	9	7	3	0
	Robot	31	5	0	4	0
Humanlike	Human	4	9	4	8	15
	Human-Robot	27	8	4	1	0
	Robot	35	3	1	1	0
Lifelike	Human	4	7	9	8	12
	Human-Robot	26	8	3	2	1
	Robot	34	4	0	2	0
Movement	Human	11	12	8	5	4
	Human-Robot	15	17	7	1	0
	Robot	19	13	6	2	0

Note. The number of ratings for each level (1 - 5) for each condition is listed. The ratings were done for fake to natural, machinelike to humanlike, artificial to lifelike, and moving rigidly to moving elegantly.

Taken together, participants evaluated the human hand to be more natural, humanlike, and lifelike; and to be moving more elegantly than the robotic hand, whereas the two conditions featuring the identical robotic hand in different agency attribution setting did not differ significantly in any of the ratings. It should be noted that movement was generally rated to be rigid rather than elegant, which is linked to the latency error of the stimulus-presentation monitor adding one frame more for each picture and thus prolonging the videos (i.e., 200 ms instead of 100 ms) and resulting in the aforementioned inter-trial-interval differences.

Fifty Percent of the Time, Tones Come Every Time: Stronger Prediction Error Effects on Neurophysiological Sensory Attenuation for Self-generated Tones

Sophie Egan[1], **Alexander Seidel**[1,2], **Constanze Weber**[1], **Marta Ghio**[1], **and Christian Bellebaum**[1]

Abstract

■ The N1/P2 amplitude reduction for self-generated tones in comparison to external tones in EEG, which has recently also been described for action observation, is an example of the so-called sensory attenuation. Whether this effect is dependent on motor-based or general predictive mechanisms is unclear. Using a paradigm, in which actions (button presses) elicited tones in only half the trials, this study examined how the processing of the tones is modulated by the prediction error in each trial in a self-performed action compared with action observation. In addition, we considered the effect of temporal predictability by adding a third condition, in which visual cues were followed by external tones in half the trials. The attenuation result patterns differed for N1 and P2 amplitudes, but neither showed an attenuation effect beyond temporal predictability. Interestingly, we found that both N1 and P2 amplitudes reflected prediction errors derived from a reinforcement learning model, in that larger errors coincided with larger amplitudes. This effect was stronger for tones following button presses compared with cued external tones, but only for self-performed and not for observed actions. Taken together, our results suggest that attenuation effects are partially driven by general predictive mechanisms irrespective of self-performed actions. However, the stronger prediction-error effects for self-generated tones suggest that distinct motor-related factors beyond temporal predictability, potentially linked to reinforcement learning, play a role in the underlying mechanisms. Further research is needed to validate these initial findings as the calculation of the prediction errors was limited by the design of the experiment. ■

INTRODUCTION

Previous research has shown that sensory stimuli resulting from self-performed actions (e.g., tones generated by self-performed button presses) are processed differently compared with identical sensory stimuli that are externally generated, which is often described as sensory attenuation (Korka, Widmann, Waszak, Darriba, & Schröger, 2022; Horváth, 2015; Hughes, Desantis, & Waszak, 2013). For example, using EEG, it was shown that the ERP components N1 and P2 are reduced in amplitude for tones that were generated by a self-performed action compared with externally generated tones (e.g., Horváth, 2015; Schafer & Marcus, 1973).

Generally, sensory attenuation is believed to be the result of predictive mechanisms (Korka et al., 2022; Horváth, 2015). It was claimed that the prediction is generated by internal forward models, indicating that these mechanisms are motor dependent in that they are based on a copy of the motor command (Wolpert & Flanagan, 2001). Alternatively, recent reviews suggested that the sensory attenuation effect for self-generated tones is due to general predictive mechanisms (Korka et al., 2022; Dogge, Custers, & Aarts, 2019). The findings concerning the role of motor processes in sensory attenuation are mixed. Some studies have shown an attenuation of N1 amplitudes for tones following own button presses even compared with externally generated tones that were temporally predictable (Klaffehn, Baess, Kunde, & Pfister, 2019; Lange, 2011; Schafer & Marcus, 1973). Others found a similar attenuation for temporally predictable tones for N1 (Harrison et al., 2021; Kaiser & Schütz-Bosbach, 2018) and P2 amplitudes (Schafer & Marcus, 1973).

A key assumption of prediction models is that the predicted sensory consequences are compared with the actual consequences of an action (or event). In case the two match, sensory attenuation occurs. In daily life, however, event-outcome contingencies are often probabilistic (e.g., when outcome-producing devices are defective, like a doorbell with low battery), leading to graded predictions and, as a consequence, to outcomes that match predictions more or less rather than in an all-or-nothing manner. In line with this, it was found that even when tones are elicited in only 50% of the trials following a button press, N1 amplitudes are attenuated relative to the processing of external tones (Seidel, Ghio, Studer, & Bellebaum, 2021; Timm, SanMiguel, Saupe, & Schröger, 2013). In other words, sensory attenuation can also occur when

---

[1]Heinrich Heine University, Faculty of Mathematics and Natural Sciences, Düsseldorf, [2]MSH Medical School Hamburg

Journal of Cognitive Neuroscience 36:10, pp. 2067–2083
https://doi.org/10.1162/jocn_a_02226

predictions are weaker, and attenuation might be strongest when the prediction is correct and there is the smallest possible prediction error. It is yet unclear to what extent sensory attenuation as reflected in N1 and P2 amplitudes is modulated by the strength of the prediction and thus by the size of the prediction error.

Recent studies directly comparing sensory attenuation with 100% and 50% action-effect contingencies (i.e., with different prediction strengths) found no difference for the N1 (Harrison, Hughes, Rudman, Christensen, & Whitford, 2023; Han, Jack, Hughes, & Whitford, 2022). For the P2, amplitudes in the 100% condition were reduced compared with amplitudes in the 50% condition (Han et al., 2022). This is in line with a study that found a reduced P2 when participants had a higher illusion of control over sound generation, which can be interpreted as the belief in a higher contingency (Seidel et al., 2021). In an alternative approach, the tone pitch (i.e., the tone identity) was altered in 30% of the trials and sensory attenuation for standard and deviant self-generated tones was compared, as the latter would also be associated with prediction errors (Knolle, Schröger, & Kotz, 2013b; see also Knolle, Schwartze, Schröger, & Kotz, 2019, for a replication with vowels generated by button press). In contrast to the studies varying tone contingency described above, the experimental manipulation affected the N1 amplitude (Knolle et al., 2013b). An attenuation of the N1 was found also for the deviant tones (for which the tone identity was mis-predicted), but the attenuation was weaker than for the standard tones, suggesting that graded predictions and prediction errors for sensory events affect the N1 (Knolle et al., 2013b).

In paradigms varying the tone contingency or identity, the tone is neither completely expected nor completely unexpected. Rather, the prediction error likely varies from trial to trial depending on the history of tone occurrences in previous trials. If the prediction, and thus the prediction error, varies along a continuum, and if sensory attenuation reflects the prediction error, this should be reflected in varying levels of sensory attenuation depending on the size of the prediction error. In the present study, we applied a paradigm in which tones followed button presses with a 50% contingency to address the question whether the size of the trial-by-trial prediction error affects the sensory attenuation effect in the processing of action-generated tones, as reflected in N1 and P2 amplitudes.

On the basis of the conception of the brain as a “prediction machine” that strives to predict incoming sensory stimuli (Clark, 2013), we hypothesized that the prediction error in studies on sensory attenuation is conceptually similar to the prediction error in feedback-learning studies, in which the error is used to update expectancies according to reinforcement learning models (Sutton & Barto, 2018; den Ouden, Kok, & de Lange, 2012). Although the elicited tones in our study do not represent a specific reward, the tone as the result of the self-performed action could be perceived as rewarding itself (Eitam, Kennedy, & Tory Higgins, 2013). Furthermore, the tone occurrence could reinforce, not the action itself, but the participant’s internal prediction of whether a tone follows the action. Such predictions are generated automatically, also in the absence of explicit instructions (Clark, 2013). If the history of tone occurrence is used as a cue to predict the next occurrence, then each tone or lack of it provides new information to predict the action outcome. We thus calculated the prediction error for each trial based on a reinforcement learning model (Rescorla & Wagner, 1972) and examined whether trial-by-trial fluctuations of the prediction error are reflected in the respective ERP components by means of a single-trial analysis. To investigate if a potential prediction error effect is specific for action-generated tones, we accounted for temporal predictability by adding a condition with cued external tones (in addition to a condition with uncued external tones), which also entailed a 50% contingency between cues and tones. Whereas attenuation studies showed inconsistent results for cued external tones (e.g., Harrison et al., 2021; Klaffehn et al., 2019), from feedback-learning studies, it is known that prediction error representations can differ depending on whether predictions are based on an action or a stimulus (Kobza & Bellebaum, 2015; O’Doherty et al., 2004).

Another open question in the literature on the processing of action consequences is whether sensory attenuation also occurs in action observation, and if so, what mechanisms are responsible for it. Because of the overlap of neural activity for self-performed action and action observation in the mirror neuron system (Rizzolatti & Sinigaglia, 2010), predictions for tones generated by the action of another human may be similar to predictions for own action-outcomes, with potential involvement of the motor system (Wolpert, Doya, & Kawato, 2003). Studies addressing sensory attenuation in action observation (Ghio, Egan, & Bellebaum, 2021; Ghio, Scharmach, & Bellebaum, 2017) found N1 attenuation (when tones were delayed relative to the eliciting action) and P2 attenuation for tones generated by an observed action, although this does not indicate that the underlying brain regions and neural mechanisms are the same as for self-performed action. For the attenuation in action observation, effects of feedback contingency and thus of prediction violation have not been examined so far. For vicarious reinforcement learning, similar processes seem to be employed as for learning from own experiences (Charpentier & O’Doherty, 2018). However, parts of the striatum as key brain structure for reinforcement learning were more involved in active learning (Kobza & Bellebaum, 2015; Bellebaum, Jokisch, Gizewski, Forsting, & Daum, 2012). To investigate the effect of prediction errors on sensory attenuation in action observation and to compare it to the one in the self-performed action, we acquired a second experimental group, in which the participants viewed picture sequences of tone-eliciting actions instead of performing the actions themselves. The conditions for the cued and uncued external tones were the same as for the performer group.

We expected that the prediction error modulates ERP amplitudes of tone processing in that a higher error yields greater amplitudes (Knolle et al., 2013b). Because of the parallels to the prediction error in reinforcement learning, we expected the thus calculated prediction errors to be better predictors of the ERP amplitudes than a prediction error based merely on the recent frequency of tone presentation. Considering the potential motor dependency of the attenuation effect (Klaffehn et al., 2019; Wolpert & Flanagan, 2001), the self-generation-specific components in the processing of mispredicted tones (Knolle et al., 2013b, 2019) and the results from feedback-learning studies on action-based prediction errors in self-performed action and action observation (e.g., Kobza & Bellebaum, 2015; Bellebaum et al., 2012), we expected the modulation-by-prediction error to be stronger for action-generated tones, especially when actions are self-performed. This effect may be especially pronounced for the P2, for which a modulation by contingency was found in a recent study (Han et al., 2022).

METHODS

Sample

In our previous studies on differences in sensory attenuation between a performer and an observer group, we recruited 20 participants per group (Ghio et al., 2017, 2021). As in the present study, we aimed to examine trial-by-trial ERP modulations by the prediction error as well as between-group differences, we recruited 80 participants, 40 per group, to increase statistical power. Seven participants were excluded as they did not fulfil the inclusion criteria, which were specified as verified right-handedness (three; reported handedness was verified with the Edinburgh Handedness Inventory; Oldfield, 1971), normal or corrected-to-normal hearing, and no history of psychiatric or neurological disorders or medication use (four). The resulting 73 participants were distributed to two groups: 36 performers (28 female participants; age: $M = 21.8$ years, $SD = 3.36$ years, range: 19–33 years) and 37 observers (28 female participants; age: $M = 23.3$ years, $SD = 4.61$ years, range: 19–37 years). Participants all gave written informed consent and received 20–25 euros or course credit for their participation. The study was approved by the ethics committee of the Faculty of Mathematics and Natural Sciences at Heinrich Heine University Düsseldorf.

Experimental Procedure

First, participants were prepared for the EEG recording and were seated in front of a 22-in. LCD computer screen with a resolution of 1680 × 1050 and a 60-Hz refresh rate. Then, they completed the training and the experiment, in which 680-Hz tones (duration 50 msec, with an increase of amplitude in the first and a decrease in the last 5 msec) were presented via Sennheiser (HD 201) headphones and responses were given via a response box (Response Pad Model RB-740 from Cedrus Corporation). For tone presentation and response recording in the training and the experiment, we used the software Presentation (Version 20.3, Build 02.25.19) from Neurobehavioral Systems. In total, the acquisition took about 2 hr.

EEG Recording

As references, we placed two electrodes on the mastoids. Via an Easycap electrode cap (www.easycap.de), further 28 passive Ag/AgCl electrode were placed at F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FC8, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, PO7, PO3, POz, PO4, and PO8 to record brain activity. An electrode at AFz served as a ground electrode. In addition, electrodes were placed above (electrode position FP2) and below the right eye, and to the sides of the eyes (positions F9 and F10) to record eye movements. BrainVision Recorder (Version 1.21.0402) and BrainAmp MR Plus amplifier (Brain Products GmbH) were used to check that impedances were kept below 20 kΩ, to record with a sampling rate of 1000 Hz (low cutoff: 0 Hz, high cutoff: 1000 Hz, software filters disabled) and to apply direct-current corrections online.

Training

The training served to familiarize participants with the 50% event-outcome association and, for performers only, to train pressing the button in a particular interval. The training comprised two parts: (1) a training for action-generated tones and (2) a training for cued tones. Only the first was different in the performer and observer groups. For performers, in the first part of the training, PsychoPy (Version 2020.2.10; Peirce et al., 2019) was used to present a sign every 2400 msec on a computer screen (40 times in total). The performers' task was to press the response button when they saw the sign. In this way, performers were trained to press a button in regular intervals, which would be their task in subsequent blocks. In 50% of the trials, a tone followed 300 msec after the button press. For the observers, the first part of the training was identical to the observer-action-tone block, in which a tone was presented 300 msec after an observed button press in 50% of the trials (for a detailed description, see below). This part of the training consisted of 40 trials. For both groups, the second part of the training was identical to the cue-tone block, in which a tone followed 300 msec after the onset of a visual cue in 50% of the trials and which is described in detail below. This part of the training also consisted of 40 trials. In each block, the instructions did not explicitly announce the 50% contingency but informed the participants that a tone may or may not follow the event.

Experiment

We included experimental blocks with tones either generated by an action, externally generated and visually cued, or externally generated without any cues, referred to as tone blocks: action, cue, and (uncued) external. To remove activity unrelated to tone processing, we included identical blocks without tones, referred to as no-tone blocks: action and cue. Each block contained 140 trials. In the tone blocks, 50% of the tones were randomly muted, resulting in 70 tone presentations per block. That means that in the action- and cue-tone blocks there were 140 actions or cues, respectively, with only half of them being followed by a tone. In the external-tone block, the 70 tones were presented without a preceding event (action or cue). The intervals between (muted) tones in the external-tone block were randomly picked from a normal distribution with a mean of 2400 msec and a standard deviation of 400 msec (approximated from previous data with a similar paradigm). This block was used for an analysis aiming at replicating the sensory attenuation effect for self-generated versus external tones found in previous studies using a 50% contingency (Seidel et al., 2021; Timm et al., 2013).

In the action blocks (action-tone and action-no-tone blocks), performers were asked to press the button in the previously learned rhythm (i.e., every 2400 msec) without cues or feedback regarding their timing. In the performer-action-tone block, a tone was presented 300 msec after the button press, and it was muted in 50% of the trials (see Figure 1A), whereas in the performer-action-no-tone-block, the tone was muted in 100% of the trials. Observers, on the other hand, viewed a picture sequence of a button press in a first-person perspective, which elicited a tone (with a delay of 300 msec) in 50% of the trials in the observer-action-tone block and no tone (i.e., muted in 100% of the trials) in the observer-action-no-tone block. The picture sequence consisted of four distinct pictures and is depicted with the duration of each picture in Figure 1B. The first picture showed the index finger placed on the response button and was presented for a variable interval (e.g., 1917 msec) to implement the randomly picked button-press interval. The intervals between the onsets of observed button presses (and thus between [muted] tones) were randomly picked from a normal distribution with a mean of 2400 msec and a standard deviation of 400 msec, which corresponds to the inter-(muted)-tone interval in the external-tone block. The second and third pictures were each presented for two frames (i.e., 33 msec) and showed the index finger

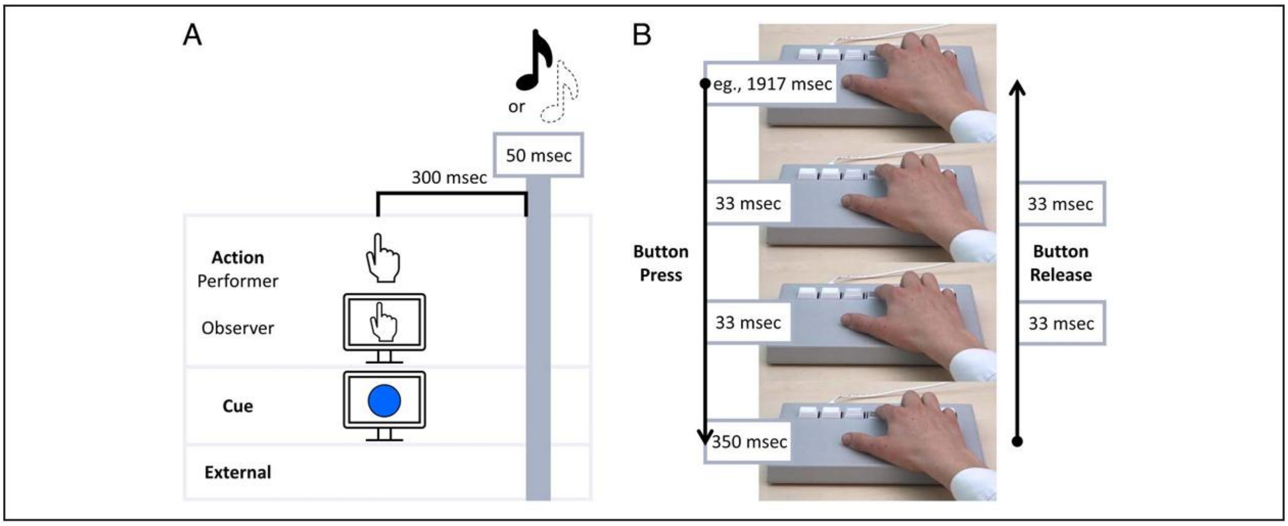


Figure 1. Experimental blocks and picture sequence. (A) The experimental conditions are depicted. The groups (performer and observer) differed only in the action condition. Tone onset was 300 msec after the onset of the button press (in the action-performer condition), of the full-button-press picture (in the action-observer condition), or of the cue (in the cue condition). The duration of the self-timed button press varied, but the durations of the observed button press and of the cue were fixed. Cue duration was 350 msec, thus covering the delay of 300 msec until tone onset and tone duration in trials with tones (for the timing of the observed button press, see B). The tones were muted (depicted in the figure with dashed outline) in 50% of the trials for the tone blocks (action, cue, and external). For the no-tone blocks (action and cue), which served to correct the ERPs for tone-unrelated potentials, the tones were muted in 100% of the trials. Apart from the muting of the tones, the trials of the no-tone blocks were identical in structure to the trials of the corresponding tone blocks. The instructions for performers in the action blocks were to press the button in the previously learned rhythm (i.e., every 2400 msec). For all other blocks, the intervals between (muted) tones were randomly picked from a normal distribution with a mean of 2400 msec and a standard deviation of 400 msec (approximated from previous data with a similar paradigm). (B) Observers viewed a picture sequence of a button press in the action condition. The relevant details of these pictures are shown. The first picture had a variable duration to implement the randomly picked interval. The other three pictures had fixed durations and followed one after the other (for button press and release) until the first picture was presented again. The picture showing the fully pressed button had a duration of 350 msec and was thus shown until the end of the tone. The pictures were either of a female (depicted) or of a male hand, matching the gender of the participant.

pressing the response button further down. The fourth picture showed the index finger pressing the response button fully down and was presented for the length of the 300-msec delay and the duration of the (muted) tone (i.e., 350 msec in total). For the button release, the second and third pictures were used in reverse order (each presented for 33 msec).

In the cue blocks, participants of both groups were presented with a blue circle on the screen, which was followed by a (muted) tone after a delay of 300 msec. The cue stayed on the screen for the length of the delay and the duration of the (muted) tone (i.e., 350 msec in total). Again, tones were muted in 50% of the trials in the tone block, and 100% of the trials in the no-tone block. The cue onset was determined in each trial to fit the randomly picked inter-(muted)-tone interval with a mean of 2400 msec and a standard deviation of 400 msec, as applied in the observer-action and external-tone blocks.

The block order was counterbalanced in that certain block combinations were presented in block sets: Half of the participants were presented with first (Block Set 1) action- and cue-tone blocks (random order), (Block Set 2) external-tone block, and then (Block Set 3) action- and cue-no-tone blocks (random order), and the other half were presented with the block sets in reversed order (3), (2), and (1). The external-tone block was placed in between to avoid carryover effects on the expectancy of the tone, when no-tone blocks directly preceded the tone blocks.

Data Analysis

The preprocessing of the EEG data was done with BrainVision Analyser (Version 2.20.7383, Brain Products GmbH). The correction of the resulting ERPs as well as the amplitude extraction and prediction error calculation described below were done with MATLAB R2019a (9.6.0.1072779, MathWorks Inc.), and the statistical analyses and visualization with R (Version 4.2.2; R Core Team, 2022).

EEG Preprocessing

First, we applied a direct current detrend, filters (Butterworth zero phase filters: low cutoff = 0.3 Hz, 12 dB/oct; high cutoff = 30 Hz, 12 dB/oct; Notch filter: 50 Hz), and an ocular correction independent component analysis (steps = 512, Infomax restricted biased). Components reflecting blinking and eye movement artifacts were selected manually and removed. Then, the data were segmented based on the tone and no-tone markers, beginning 200 msec before with a total length of 800 msec. Only trials meeting the response time criteria were considered (the interval to the preceding button press was between 1800 and 10,000 msec, and the button press was not part of a double press, i.e., button presses were excluded if a second button press followed less than 800 msec after it). The first 200 msec of the segment was used for baseline correction. We applied the automatic procedure in BrainVision Analyser to discard segments with artifacts (maximal allowed voltage step of 50 μV/msec, maximal allowed difference of values of 100 μV in 100-msec intervals, allowed amplitudes between −100 and 100 μV, and lowest allowed activity of 0.5 μV in 100 msec). The main analysis was based on the single-trial ERPs of the tone blocks, which entailed action-tone-, cue-tone- and external tone-related ERPs. For action-tone and cue-tone single-trial ERPs, corrections were performed to account for ERPs related to the actions or cues themselves. For this purpose, ERP averages of no-tone blocks were calculated (i.e., for the action-no-tone and the cue-no-tone). The differences of action-tone single-trial ERPs and the averaged action-no-tone ERPs made up the single-trial ERPs of the action condition. The differences of cue-tone single-trial ERPs and the averaged cue-no-tone ERPs made up the single-trial ERPs of the cue condition. The external-tone single-trial ERPs (i.e., the external condition) were not corrected, as the tone was not preceded by any event.

Single-trial Amplitudes

The electrodes Fz, FC3, FCz, FC4, and Cz were considered for analysis. We chose this frontocentral electrode cluster, because the N1 and P2 topographies (see Figure 2) showed that both components were most pronounced at frontocentral electrode sites, including three midline electrodes, which were considered in previous studies (Klaffehn et al., 2019; Kaiser & Schütz-Bosbach, 2018; Ghio et al., 2017). In accordance with a procedure for single-trial analyses we applied in Seidel, Weber, Ghio, and Bellebaum (2023) and Albrecht and Bellebaum (2021), participant averages were used to determine the time windows appropriate for the extraction of mean amplitudes of the ERP components of interest in each trial. Thus, ERPs were averaged for each participant; for the action, cue, and external condition; and for the specified electrodes. Then, in a first step, peaks were defined for the N1 component as the local minimum in the interval between 50 and 150 msec after tone onset, and for the P2 component as the local maximum in the interval between 135 and 300 msec after tone onset for each participant in each condition at each electrode (see Figure 2). In a second step, single-trial amplitudes for the N1 and P2 components were defined as the mean amplitude in the 40-msec interval around the respective peak latency for each participant, condition, and electrode.

Prediction Error Calculation

As outlined in the Introduction section, we aimed to calculate a prediction error for each trial, which represented

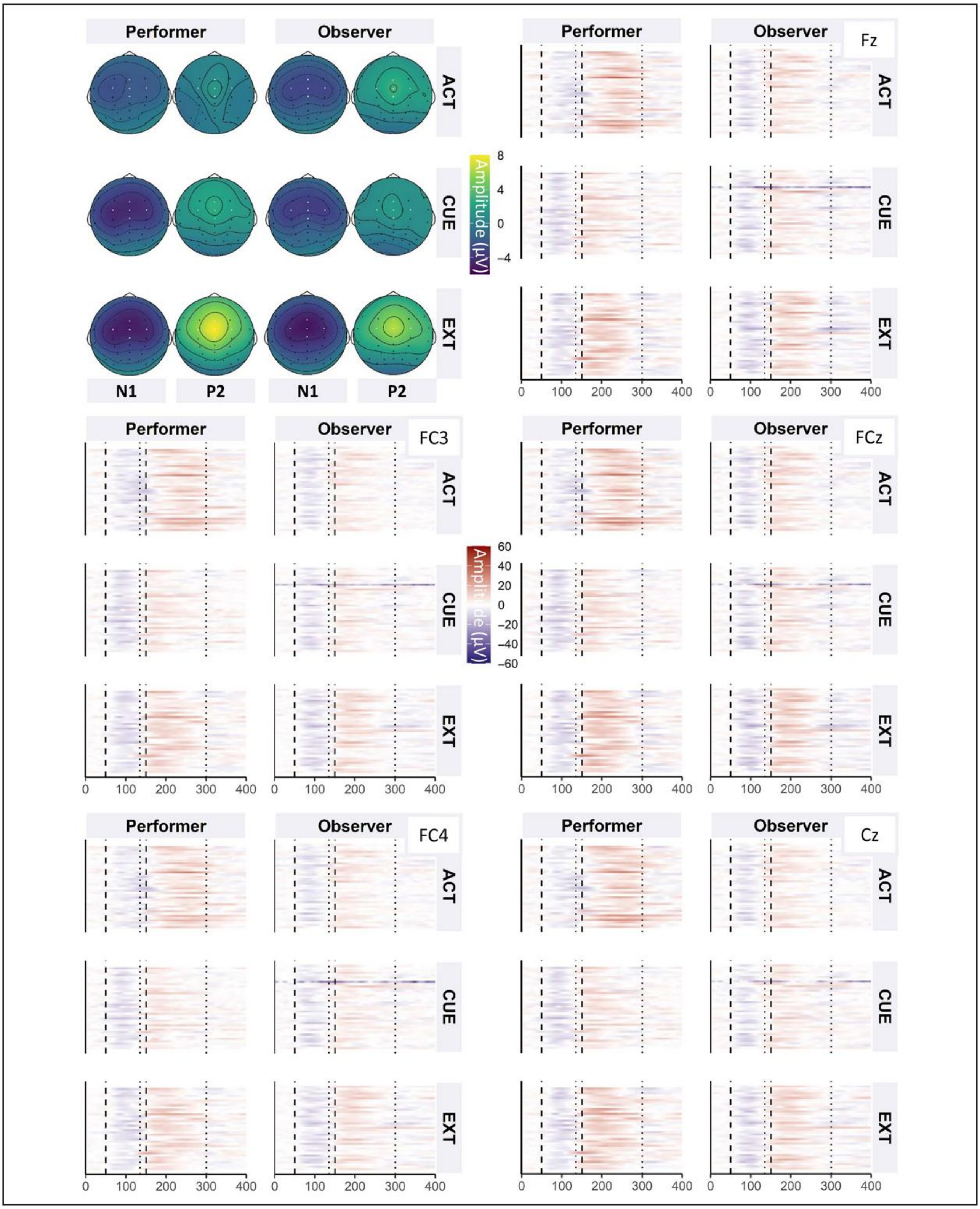


Figure 2. ERPs and topographies. For the topographies, ERPs were aligned to the 40-msec interval as used for N1 and P2 amplitudes, respectively, and the average amplitudes of that interval are presented for each group, condition, and component. The displayed ERPs were averaged across trials for each participant, condition, and electrode. For the action and cue conditions, condition refers to the difference of tone single-trial ERPs and averaged no-tone ERPs. The amplitude is color-coded (white-centered), and the participant averages are stacked on top of each other separately for the performer and observer groups. To display the N1 and P2 time windows, only the 400 msec after the tone presentation (at 0 msec) are depicted with dashed lines enclosing the interval for N1 peak detection and dotted lines enclosing the interval for P2 peak detection.

how surprising the occurrence of the tone was. To do so, we used the Rescorla–Wagner model known from reinforcement learning (Zhang, Lengersdorff, Mikus, Glascher, & Lamm, 2020; Rescorla & Wagner, 1972), in which the subjective value of a stimulus is updated based on prediction errors and a specific learning rate. This seems unusual, as participants in our task were not asked to learn to predict whether an action would be followed by a tone. According to accumulating evidence that our brain continuously makes predictions concerning incoming sensory stimuli (Clark, 2013), however, it is likely that participants made sensory-stimuli-related predictions, even if they were not instructed to do so. The subjective expectancy of tone occurrence would then correspond to the subjective reward expectancy in the reinforcement learning model. The prediction error was calculated in every trial as the difference between the event that occurred and the expectancy. The event was either the tone occurrence (coded as 1) or the omission of the tone (coded as 0). The expectancy was initialized as 0.5 and updated each trial by adding the prediction error weighted with the learning rate. As we did not have any behavioral measure concerning participants' expectancy, we used fixed learning rates to update it from trial to trial and calculate the prediction error similarly to O'Doherty, Dayan, Friston, Critchley, and Dolan (2003). We tested two different learning rates, a low (i.e., 0.2) and a high (i.e., 0.7) one, corresponding to slight or strong updating of expectancy from trial to trial (cf. O'Doherty et al., 2003).

In an alternative approach, we calculated the expectancy based on the frequency of tone occurrences in the last 10 trials preceding the current trial. Hence, the expectancy was not updated with a weighted prediction error. The prediction error was then calculated as the difference between the event and the expectancy.

Taken together, we had single-trial values of three types of prediction errors: based on a low and high learning rate, and on frequency. With this procedure, we could test if prediction errors based on reinforcement learning models are more strongly related to ERP amplitudes following tones than prediction errors based on frequency. Note that, because only the tone trials were further analyzed, the possible range of the prediction error value was between 0 and 1 for all statistical analyses.

Statistical Analyses

Two independent statistical analyses were conducted. For both analyses, we fitted mixed linear models and included the factors tone trial type (within-subject) and group (between-subject; performers and observers). The first analysis aimed to examine whether a pattern of sensory attenuation occurs in our paradigm with 50% contingency between preceding event (action or cue) and tone. In this analysis, the factor tone trial type consisted of three levels: external, cue, and action (external as reference). We calculated a model with both factors with restricted maximum likelihood and the optimizer *bobyca* using the package *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017). This analysis was done separately for N1 and P2 amplitudes. Random effects based on participant and electrode were manually reduced starting from the maximal model till the model converged, leaving random intercept and slope for the tone conditions by participant and random intercept by electrode for both components. Thus, the model for N1 and P2 was: Amplitude ~ tone trial type * group + (1 + tone trial type | participant) + (1 | electrode). We then applied the ANOVA function as implemented in the *lmerTest* package for the decomposition of fixed-effects contributions (the effect size was calculated with the package *effectsize*; Ben-Shachar, Lüdecke, & Makowski, 2020) and resolved the significant interaction effects with Bonferroni-corrected pairwise comparisons using the package *emmeans* (degrees-of-freedom method: Satterthwaite; Lenth, 2022).

In the second analysis, which was the one of main interest, aiming to examine a potential modulation of N1 and P2 amplitudes by the prediction error, we considered only conditions in which an event preceded the tone, namely, the action and cue conditions. The factor tone trial type thus consisted of two instead of three levels. The prediction error was introduced as an additional factor, which was standardized (taking the reference point from no error to the mean error). We set up three models, one for each type of prediction error. For the frequency prediction error, the first 10 trials were not considered in the analysis. We manually reduced the random effects till the same model for each prediction error type converged, leaving for the N1 and P2 random intercept and slopes for the tone conditions, prediction error, and their interaction by participant; random intercept and slope for the tone conditions by electrode for the N1; and only random intercept by electrode for the P2. The models were therefore: N1 Amplitude ~ tone trial type * group * prediction error + (1 + tone trial type * prediction error | participant) + (1 + tone trial type | electrode) and P2 Amplitude ~ tone trial type * group * prediction error + (1 + tone trial type * prediction error | participant) + (1 | electrode). To determine which type of prediction error yielded the best model fit, we compared the models of the different prediction errors that were calculated on the same data set (i.e., without the first 10 trials) and with maximum likelihood instead of restricted maximum likelihood (recommended by Meteyard & Davies, 2020). The package *performance* (Lüdecke, Ben-Shachar, Patil, Waggoner, & Makowski, 2021) was used for the comparison via Bayes Factors and other performance indices (e.g., the Akaike information criterion). We recalculated the best-fitting model with restricted maximum likelihood and, for learning rate prediction errors, on the original data set (i.e., including the first 10 trials). As before, we then calculated an ANOVA and resolved significant interaction effects by calculating pairwise comparisons or separate models.

RESULTS

Sensory Attenuation with 50% Contingency

Figure 3 displays the averaged ERPs. For the N1, the main effect of Tone Trial Type was significant, $F(2, 71.2) = 11.41, p < .001, \eta_p^2 = .24$. Amplitudes for action tones were significantly reduced compared with external tones, but not compared with cue tones, whereas amplitudes for cue tones and external tones did not differ significantly (see Figure 4 for descriptive statistics and statistical indices). The main effect of Group, $F(1, 71.0) = 0.18$, $p = .67, \eta_p^2 < .01$, and the interaction of Tone Trial Type and Group, $F(2, 71.2) = 0.64, p = .53, \eta_p^2 = .02$, were not significant.

For the P2, all effects were significant. The main effect of Tone Trial Type, $F(2, 71.2) = 44.61, p < .001, \eta_p^2 = .56$, and the main effect of Group, $F(1, 71.0) = 10.82, p = .002, \eta_p^2 = .13$, were further moderated by a significant interaction, $F(2, 71.2) = 7.80, p < .001, \eta_p^2 = .18$, which was resolved by separate pairwise comparisons for the groups. As can be seen in Figure 4, which shows descriptive statistics and statistical indices, amplitudes for cue tones were attenuated compared with external tones in both groups.

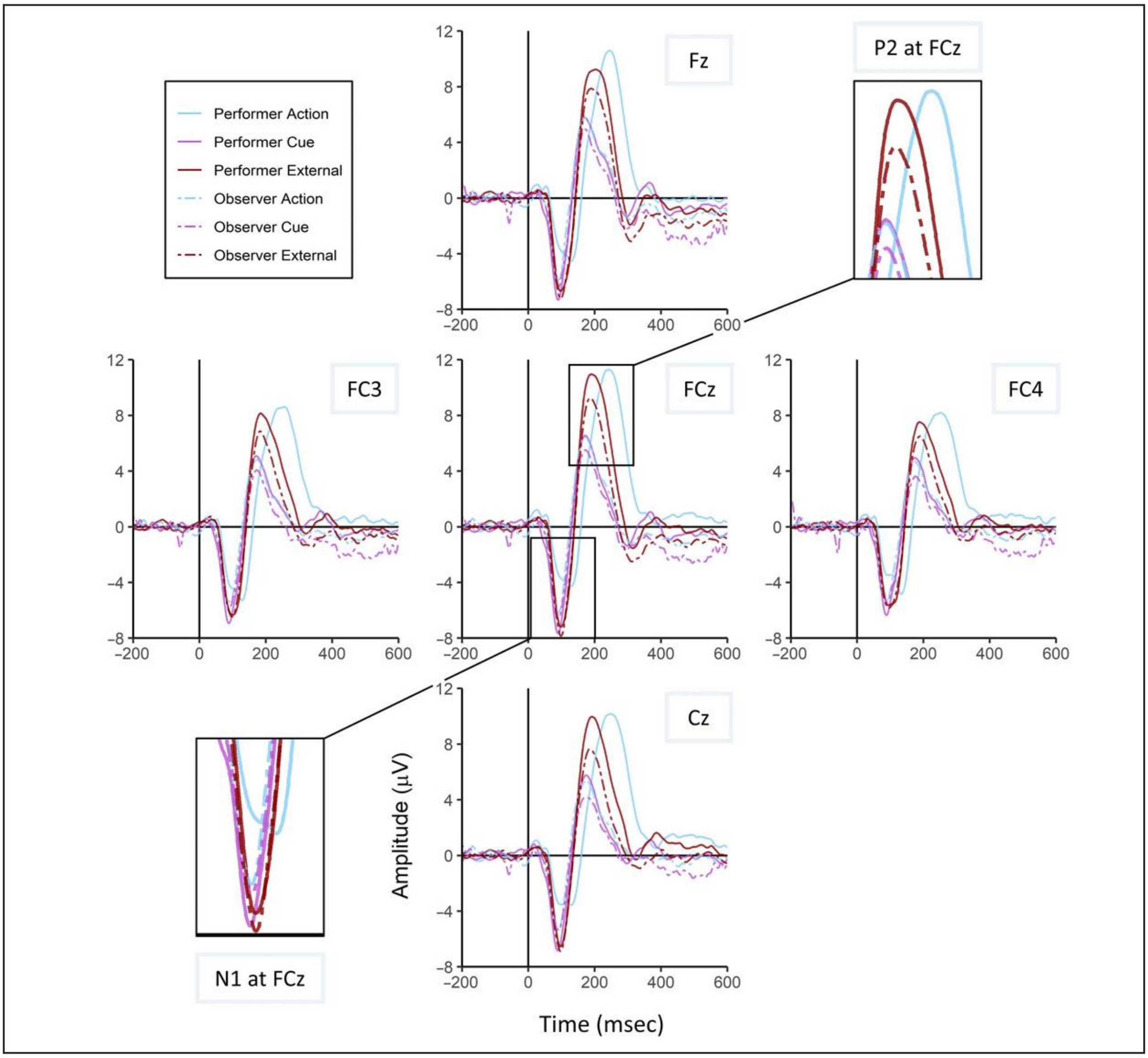


Figure 3. Averaged ERPs. The displayed ERPs were averaged across trials and participants for each group, condition, and electrode. For the action and cue conditions, condition refers to the difference of tone single-trial ERPs and averaged no-tone ERPs. Because these are averages across participants, the ERPs do not directly correspond to our statistical analyses of the N1 and P2 amplitudes, for which the latencies for amplitude extraction were determined by participant averages.

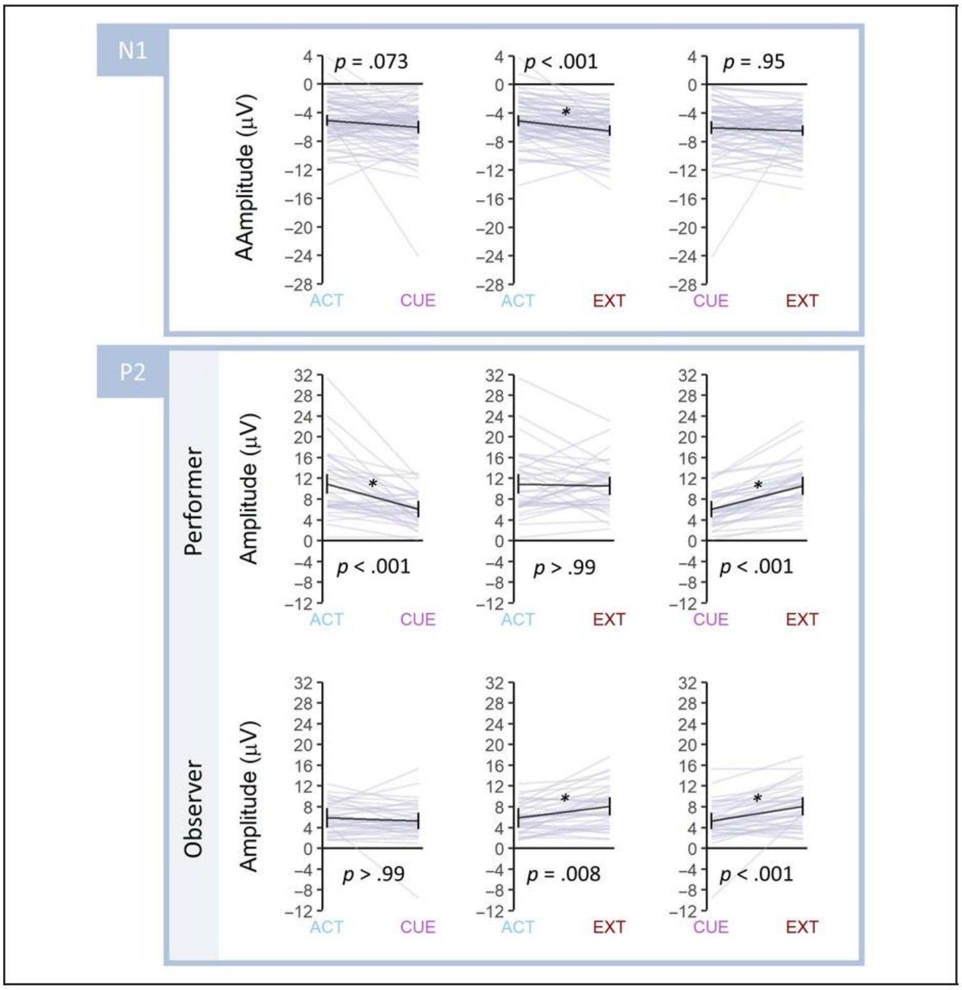


Figure 4. Pairwise comparisons for the sensory attenuation effect. For the N1, the pairwise comparisons for the significant main effect of tone trial type (action [ACT], cue [CUE], and external [EXT]) are presented. For the P2, because of the significant interaction with group, the pairwise comparisons for the effect of tone trial type are shown separately for performers and observers. The means across single trials and electrodes for each participant are depicted in gray; the estimated means across participants with 95% confidence intervals are in black.

In performers, amplitudes for action tones were not attenuated but rather comparable to those for external tones and even enhanced relative to cue tones. In observers, in turn, action tones elicited comparable amplitudes as cue tones and attenuated amplitudes relative to external tones.

Prediction Error Modulation

For the N1, the best-fitting model was provided by the prediction error calculation based on the low learning rate (see Table 1). For this model, the main effects of Tone Trial Type (including only action and cue), $F(1, 68.9) = 5.01, p = .028, \eta_p^2 = .07$, and of Prediction Error, $F(1, 70.5) = 38.79, p < .001, \eta_p^2 = .36$, were significant. More importantly, the Tone Trial Type, Prediction Error, and Group interaction was significant, $F(1, 69.3) = 6.68, p = .012, \eta_p^2 = .09$, whereas no other effects were, all $p > .064$. We resolved the interaction by calculating separate models (see Table 2 for fixed effects and model performance indices of all subsequent models) for performers and observers including the same random effects as before: Amplitude ~ tone trial type * prediction error + (1 + tone trial type * prediction error | participant) + (1 + tone trial type | electrode). In the performer group, all effects were significant, Tone Trial Type: $F(1, 34.6) = 4.49, p = .041, \eta_p^2 = .11$; Prediction Error: $F(1, 34.8) = 22.83, p < .001, \eta_p^2 = .40$; interaction: $F(1, 33.3) = 9.67, p = .004, \eta_p^2 = .23$, and separate models for the tone trial types were calculated including only random intercept and slope for prediction error by participant and random intercept by electrode: Amplitude ~ prediction error + (1 + prediction error | participant) + (1 | electrode). Although amplitudes for both action and cue tones were significantly modulated by prediction error in that higher prediction errors coincided with larger (i.e., more negative) N1 amplitudes, this was stronger for action tones (see Figure 5). For observers, only the Prediction Error main effect was significant, $F(1, 35.6) = 16.87, p < .001, \eta_p^2 = .32$ (Tone Trial Type and interaction $p > .27$), and thus amplitudes for action and cue tones were

Table 1. Prediction Error Mixed Linear Models Summaries and Comparison

	N1			*P2*		
	LLR	*HLR*	*FQ*	*LLR*	*HLR*	*FQ*
Fixed effects:						
Intercept	−6.428	−6.455	−6.454	6.130	6.133	6.128
TTT action	1.176	1.196	1.196	4.590	4.644	4.591
PE	−0.638	−0.716	−0.446	0.993	1.136	0.767
Group obs	0.702	0.728	0.732	−0.810	−0.826	−0.829
TTT action : PE	−0.771	−0.411	−0.699	1.233	1.226	1.052
TTT action : Group obs	−0.474	−0.474	−0.494	−4.061	−4.084	−4.030
PE : Group obs	0.019	0.090	0.127	−0.366	−0.295	−0.388
TTT action : PE : Group obs	0.761	0.521	0.538	−1.123	−0.996	−0.952
AIC	3.094e+5	3.096e+5	3.097e+5	3.224e+5	3.221e+5	3.228e+5
BIC	3.096e+5	3.098e+5	3.099e+5	3.226e+5	3.223e+5	3.230e+5
Cond. R2	0.157	0.151	0.150	0.216	0.219	0.208
Marg. R2	0.013	0.011	0.008	0.051	0.055	0.044
ICC	0.147	0.142	0.143	0.174	0.174	0.171
BF comparison to LLR		1.22e−48	1.41e−76		7.03e+60	4.20e−98

To analyze the potential prediction error (PE) modulation, mixed linear models with tone trial type (TTT; cue and action), PE, and group (performers and observers [obs]) were specified. They were calculated with each type of PE (low learning rate [LLR], high learning rate [HLR], and frequency [FQ]) for the N1 and P2 amplitudes on the same data set for the comparison (i.e., without the first 10 trials excluded for the FQ PE). Specifically, the models were: N1 Amplitude ~ TTT * group * PE + (1 + TTT * PE | participant) + (TTT | electrode) and P2 Amplitude ~ TTT * group * PE + (1 + TTT * PE | participant) + (1 | electrode). The "*" includes all interactions and main effects; the ":" refers only to the specified interaction. The number of data points for each model was 43190. AIC = Akaike information criterion; BF = Bayes Factor; BIC = Bayesian information criterion; Cond. = Conditional; ICC = intraclass correlation coefficient; Marg. = Marginal.

Table 2. Model Summaries for Interaction Resolutions

	Fixed Effects	*Estimate*	*SE*	*df*	*t*	*p*
N1						
Performer	Intercept	−6.395	0.462	35.0	−13.83	<.001
	TTT action	1.216	0.574	34.6	2.12	.041
	PE	−0.630	0.220	32.9	−2.87	.007
	TTT : PE	−0.762	0.245	33.3	−3.11	.004
Action	Intercept	−5.175	0.554	36.3	−9.34	<.001
	PE	−1.393	0.268	34.4	−5.20	<.001
Cue	Intercept	−6.397	0.463	35.1	−13.82	<.001
	PE	−0.643	0.224	31.7	−2.88	.007

Table 2. *(continued)*

	Fixed Effects	*Estimate*	*SE*	*df*	*t*	*p*
Observer	Intercept	−5.715	0.744	37.5	−7.68	<.001
	TTT action	0.728	0.659	36.1	1.10	.277
	PE	−0.694	0.164	45.5	−4.23	<.001
	TTT : PE	0.120	0.113	22937.0	1.07	.287
P2						
Performer	Intercept	6.040	0.710	20.8	8.50	<.001
	TTT action	4.961	0.934	35.1	5.31	<.001
	PE	1.180	0.226	35.1	5.21	<.001
	TTT : PE	1.312	0.455	35.1	2.88	.007
Action	Intercept	11.001	1.185	32.5	9.28	<.001
	PE	2.492	0.369	34.9	6.75	<.001
Cue	Intercept	6.041	0.633	32.1	9.55	<.001
	PE	1.179	0.227	35.1	5.20	<.001
Observer	Intercept	5.281	0.747	31.6	7.07	<.001
	TTT action	0.589	0.557	36.1	1.06	.298
	PE	0.875	0.242	35.7	3.62	<.001
	TTT : PE	0.228	0.260	35.3	0.88	.386

			R2		
	AIC	*BIC*	*Cond.*	*Marg.*	*ICC*
N1					
Performer	1.654e+5	1.656e+5	0.134	0.018	0.119
Action	8.388e+4	8.393e+4	0.161	0.022	0.142
Cue	8.157e+4	8.162e+4	0.099	0.005	0.094
Observer	1.647e+5	1.648e+5	0.176	0.006	0.171
P2					
Performer	1.735e+5	1.736e+5	0.262	0.066	0.210
Action	8.961e+4	8.966e+4	0.289	0.036	0.263
Cue	8.364e+4	8.369e+4	0.134	0.013	0.123
Observer	1.704e+5	1.705e+5	0.139	0.010	0.131

To resolve the three-way interactions for the N1 model (PE based on LLR) and the P2 model (PE based on HLR), separate mixed linear models were calculated for performers and observers. For both, the TTT (cue and action) and PE interaction was significant in the performer group, and separate models for the TTTs were calculated. The ":" refers only to the specified interaction. Number of data points were 22,965 for the performer group, 11,640 for performer-action, 11,325 for performer-cue, and 23,135 for the observer group. AIC = Akaike information criterion; BIC = Bayesian information criterion; Cond. = Conditional; ICC = intraclass correlation coefficient; Marg. = Marginal.

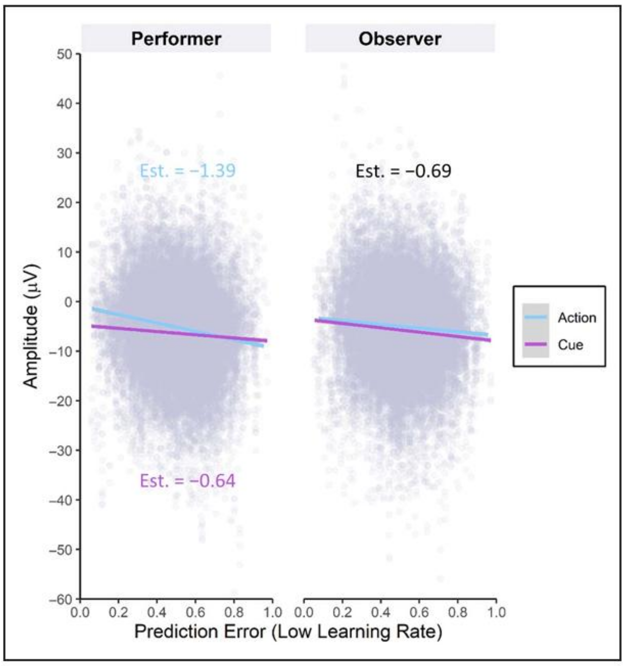


Figure 5. N1 amplitudes and prediction error. The prediction error calculated based on the reinforcement model with the low learning rate was analyzed for the N1. To resolve the three-way interaction, separate mixed linear models were calculated for performers and observers. The tone trial type (action and cue) and prediction error interaction was significant for performers but not for observers. Therefore, the prediction error effect was calculated separately for action and cue tones only for performers. The fixed effect estimates (Est.) of the respective main effect of prediction error are added. The depicted lines may vary slightly from the modeled effects because random effects were not considered.

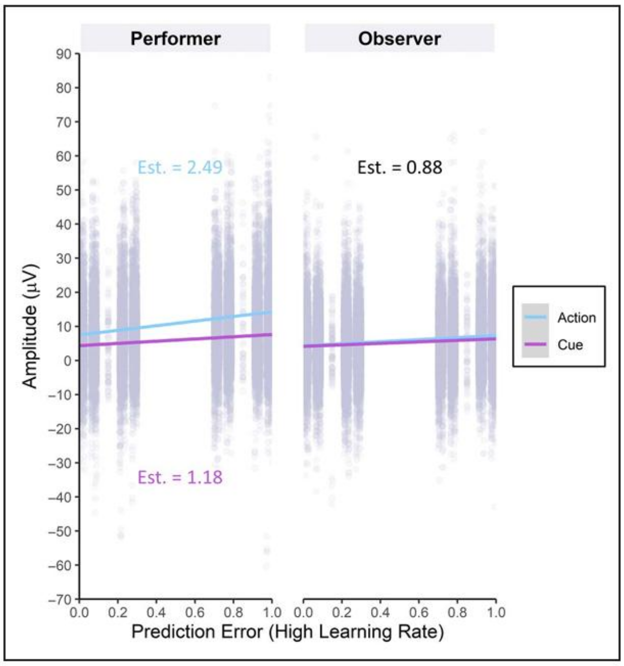


Figure 6. P2 amplitudes and prediction error. The prediction error calculated based on the reinforcement model with a high learning rate was analyzed for the P2. To resolve the three-way interaction, separate mixed linear models were calculated for performers and observers. The tone trial type (action and cue) and prediction error interaction was significant for performers but not for observers. Therefore, the prediction error effect was calculated separately for action and cue tones only for performers. The fixed effect estimates (Est.) of the respective main effect of prediction error are added. The depicted lines may vary slightly from the modeled effects because random effects were not considered.

similarly related to the prediction error (see Figure 5). Again, higher prediction errors led to larger (i.e., more negative) N1 amplitudes.

For the P2, the best-fitting model was reached with the prediction error calculation based on the high learning rate (see Table 1 for the comparison). All main effects, Tone Trial Type: $F(1, 71.2) = 26.39, p < .001, \eta_p^2 = .27$; Prediction Error: $F(1, 71.3) = 111.91, p < .001, \eta_p^2 = .61$; Group: $F(1, 71.1) = 12.03, p < .001, \eta_p^2 = .14$; and interactions, Tone Trial Type and Prediction Error: $F(1, 71.0) = 8.71, p = .004, \eta_p^2 = .11$; Tone Trial Type and Group: $F(1, 71.2) = 16.31, p < .001, \eta_p^2 = .19$; Prediction Error and Group: $F(1, 71.3) = 10.15, p = .002, \eta_p^2 = .12$, were significant, the three-way interaction being of particular interest, $F(1, 71.0) = 4.38, p = .040, \eta_p^2 = .06$. As for the N1, we calculated separate models (see Table 2 for fixed effects and model performance indices of all subsequent models) for performers and observers including the same random effects: Amplitude ~ tone trial type * prediction error + (1 + tone trial type * prediction error | participant) + (1 | electrode). For performers, all effects were significant, tone trial type: $F(1, 35.1) = 28.23, p < .001, \eta_p^2 = .45$; prediction error: $F(1, 35.3) = 79.99, p < .001, \eta_p^2 = .69$; interaction: $F(1, 35.1) = 8.32, p = .007, \eta_p^2 = .19$, and separate models for the tone trial types were calculated, now including only random intercept and slope for prediction error by participant and random intercept by electrode: Amplitude ~ prediction error + (1 + prediction error | participant) + (1 | electrode). Again, amplitudes for action tones were more strongly related to prediction errors than those for cue tones (see Figure 6) and higher prediction errors led to larger (i.e., more positive) amplitudes. The same pattern as for the N1 was also evident for the observer group, in that only the prediction error main effect was significant, $F(1, 35.8) = 33.52, p < .001, \eta_p^2 = .48$ (all other $p > .29$), and amplitudes for action and cue tones were larger with higher prediction errors (see Figure 6).

DISCUSSION

In this study, we first investigated the attenuation of N1 and P2 amplitudes in a 50%-contingency paradigm for action performance and action observation, considering also effects of temporal predictability via visually cued external tones. Then, in our analysis of main interest, we examined the effects of expectancy on the attenuation of N1 and P2 amplitudes by analyzing the modulation of single-trial amplitudes by the prediction error. We found significant N1 attenuation relative to uncued external tones only for action tones (i.e., generated by a performed or an observed action) but not for cued tones. For the P2, the performers showed (late) amplitudes for action tones that were not attenuated compared with uncued tones. However, we found P2 attenuation compared with uncued tones for tones generated by an observed action and for cued tones. Interestingly, despite the differences in N1 and P2 amplitude attenuation results, the pattern of results for the modulation-by-prediction error was the same. Higher prediction errors were significantly related to higher amplitudes. This effect was stronger for tones following actions than for tones that were cued by a visual stimulus, but only when actions were self-performed.

The N1 Component

We replicated the common finding of attenuated N1 amplitudes of self-generated tones compared with uncued external tones (Horváth, 2015; Schafer & Marcus, 1973) even with only a 50% contingency (Seidel et al., 2021; Timm et al., 2013), and further found no significant difference between N1 amplitudes of tones generated by self-performed and observed actions. This is in line with previous observation studies on sensory attenuation applying a delay between observed action and tone (Egan, Ghio, & Bellebaum, 2023; Ghio et al., 2017), similar to the present study. This could imply action-specific processes underlying the N1 attenuation that apply for an action performance as well as observation (Wolpert et al., 2003). However, temporal predictability remains an important factor as self-performed actions allow for a highly reliable temporal prediction of the outcome (if the temporal relationship of action and outcome is fixed) and we ensured temporal predictability of the observed actions' outcome via the action-outcome delay of 300 msec. In an observation study without an action-outcome delay, for instance, there was no N1 attenuation for observers (Ghio et al., 2021). To account for the effect of temporal predictability in this study, we included external tones that were visually cued. The N1 amplitudes in the cue condition were not reduced compared with those of external tones, which might suggest that N1 attenuation is indeed action specific, as previous studies showed (e.g., Klaffehn et al., 2019). However, we cannot draw firm conclusions as the N1 amplitudes in the cue condition also did not differ significantly from the amplitudes in the action condition, in line with studies providing evidence that N1 attenuation is not action specific but rather reflects more general prediction mechanisms (Egan et al., 2023; Harrison et al., 2021; Kaiser & Schütz-Bosbach, 2018).

Interestingly, we found a significant modulation-by-prediction error for action-generated (self-performed and observed) and cued tones: The more expected the tones, the smaller the N1 amplitudes, which is in line with findings for deviant tones (Knolle et al., 2013b). However, there are differences between deviant tones in previous studies and the unexpected tones in our study to consider. Altering the tone identity in a small percentage of the trials enabled the participants to form precise predictions, making the deviant tones mispredicted rather than unpredicted tones (Hsu, Hämäläinen, & Waszak, 2018). For such mispredicted tones, sensory attenuation for action-generated versus external tones was still observed (Knolle et al., 2013b), although the tone likely elicited a larger prediction error than the more frequently presented standard tone. Yet, because the predictions in deviant-tone studies refer to tone identity, that is, to the "what dimension" in stimulus prediction, the results cannot directly be compared with the results of the present study, which dealt with predictions concerning "whether" a stimulus would occur or not.

Still, the significant modulation-by-prediction error for both action-generated and cued tones indicates that predictive mechanisms influenced the N1 amplitudes (Dogge et al., 2019; Horváth, 2015) and participants likely formed predictions regarding the occurrence of the tone on each trial. Furthermore, the significant three-way interaction revealed a stronger prediction-error effect for self-generated tones. This is in contrast to our N1 attenuation results showing no difference between action performance and observation but would be in line with the proposed action-performance specificity in deviant processing (Knolle et al., 2013b, 2019). The stronger effect for self-generation might be because of the higher reliability and specificity of the prediction or the greater saliency for fine-tuning the predictive mechanisms regarding one's own actions (Knolle et al., 2013b). This finding might support the internal forward models and the dependency on motor-related processes (i.e., the efference copy) in forming the predictions (Wolpert & Flanagan, 2001).

For the analyses of the effects of variable expectancies from trial to trial on sensory attenuation, we considered three types of prediction errors and found the best fitting for the N1 to be the prediction error based on a reinforcement learning model with a low learning rate, as opposed to one based on a high learning rate or on the statistics of tone occurrences. Although they were not instructed to do so, participants may have predicted the tone occurrence after their action and updated their expectancy similarly as they would to learn response-outcome contingencies to maximize the reward. This would support the notion of a general underlying computation of prediction errors across different domains (Sutton & Barto, 2018; den Ouden et al., 2012) and would be in line with error-driven learning accounts for many different types of associations including those involved in language acquisition (Nixon, 2020). According to these accounts, reinforcement learning mechanisms require a cue-outcome structure. In our experimental paradigm, the (observed) action or the visual stimulus (i.e., blue circle) provided the cue, based on which the outcome (i.e., tone) could be predicted with varying certainty from trial to trial. Studies from the language domain showed that the successful prediction of a new word from context recruits the striatum to a similar extent as the processing of rewards (Ripollés et al., 2014). Along similar lines, striatal activations reflect trial by trial predictions derived from a reinforcement learning model in learning syntactic structures (Orpella, Mas-Herrero, Ripollés, Marco-Pallarés, & de Diego-Balaguer, 2021). These findings suggest that the processing of fulfilled and not-fulfilled predictions recruits neural structures of the reward system, even if no explicit rewards are involved, which may also underly the N1 and P2 amplitude modulations by prediction errors derived from a reinforcement learning model. Whether and to what extent structures of the reward system are recruited in predicting sensory outcomes needs to be addressed in future studies.

The stronger prediction error effect for tones following own actions than for tones following observed actions and cued tones can also be interpreted in the reinforcement learning framework: It has been shown that reward prediction error representations in relevant brain areas, such as the striatum and cerebellum, are stronger for rewards following own than observed actions (Kobza & Bellebaum, 2015; Bellebaum et al., 2012). The low learning rate led to a slight expectancy increase or decrease from trial to trial, as is typical for nondeclarative learning including feedback learning (Shohamy, Myers, Kalanithi, & Gluck, 2008). The slow change might reflect that the N1, as an early ERP component, is related to more automatic tone processing (e.g., Knolle, Schröger, & Kotz, 2013a).

As a first note of caution, it has to be mentioned, however, that the prediction error calculation was based on arbitrarily fixed learning rates across all participants, as we did not have a behavioral measure reflecting the actual updating of participants' predictions. Although a similar procedure has been applied for the modeling of learning in a classical conditioning task (O'Doherty et al., 2003), this limits the interpretation of our finding. Furthermore, it should be noted that the N1 has been shown to be affected by intertone intervals, with shorter intervals leading to smaller amplitudes (Sanmiguel, Todd, & Schröger, 2013), and that in our prediction error calculation, this is invariably contained as, for example, the omission of tones will lead to less expectancy and to longer intertone intervals. Indeed, all our prediction error types correlated significantly with the intertone intervals (all $p < .001$). Although this may explain the prediction error main effect, it cannot explain the stronger prediction error modulation of N1 amplitudes for self-generated tones. Taken together, our results imply that different underlying processes may lead to similar overall attenuation patterns. The specifically strong modulation by the prediction error for tones following own actions resembles findings from reinforcement learning where (reward) outcomes of own actions are also processed differently compared with outcomes not following own actions (Kobza & Bellebaum, 2015; Bellebaum et al., 2012; Bellebaum, Kobza, Thiele, & Daum, 2010; Yeung, Holroyd, & Cohen, 2005).

The P2 Component

For the P2, we found attenuation for temporally predictable tones by cueing or action observation via picture sequence compared with uncued external tones. Thus, it seems that action observation provides no motor-related advantage over any other visual cueing in the attenuation of the P2 amplitude. Regarding the attenuation for action observation, our results are in line with previous studies using a picture sequence (Seidel et al., 2023; Ghio et al., 2017), and studies using live action observation (Egan et al., 2023; Ghio et al., 2021). However, Egan and colleagues (2023) included a cue condition and, although finding P2 attenuation for cued compared with uncued external tones, the attenuation was stronger for tones generated by the other participant (i.e., action observation), whereas Harrison and colleagues (2021), for instance, did not find P2 attenuation for cued external tones at all. Given the significant difference in P2 amplitudes between action performance and observation in the previous picture-sequence-observation study (Ghio et al., 2017) and the lack of one in the previous live-observation studies (Egan et al., 2023; Ghio et al., 2021), we assume that the setting of the action observation (live or via picture sequence) may affect the underlying processes.

For action performance (i.e., self-generation), however, we found late and unattenuated P2 amplitudes, similar to Seidel and colleagues (2021), who also applied a paradigm with the action-tone contingency reduced to 50%. Whereas the temporal onset of tones was at least as predictable as in the cue and observation conditions, the P2 has also been related to sense or judgment of agency (Timm, Schönwiesner, Schröger, & SanMiguel, 2016) and the belief in contingency (Seidel et al., 2021). Although we did not acquire agency judgments, it is possible that through the randomness of the occurrence of the sensory outcome, participants felt less agency for the tone and so the P2 was not attenuated. Alternatively, as Seidel and colleagues (2021) discussed, there could be a P3 overlay causing the increased latencies and amplitudes. A later positivity (i.e., P3 or P3a) has been reported for infrequent tones and interpreted in terms of attention orientation to the tone (Knolle et al., 2013a, 2013b). The uncertainty of the outcome occurrence might thus have led to increased attention to it (similar to the attention afforded to an unpredictable external tone) but only selectively for own action outcomes, perhaps because of the greater saliency for oneself (Knolle et al., 2013b).

Despite the different pattern of results in the attenuation analysis for the N1 and P2, we found comparable prediction error effects on P2 amplitudes as on the N1 amplitudes. The more expected the tone, the smaller the P2 amplitudes, as was found by Han and colleagues

(2022), and this effect was stronger in action performance. However, we found that, rather than a reinforcement learning model with a low learning rate as for the N1, a model with a high learning rate (i.e., greater impact of prediction errors on subsequent expectancies) fitted the data best. Hence, we found no clear dissociation of the components with regard to the prediction error effects but yet an indication that the processes reflected by the N1 and P2 differ, like many studies investigating sensory attenuation have reported (e.g., Knolle et al., 2013a). Previous studies that described differences in the N1 and the P2 patterns suggested a more conscious process underlying the P2 attenuation, such as a judgment of agency (Timm et al., 2016). Considering the reduction of P2 amplitudes when the contingency was merely believed to be higher (Seidel et al., 2021), the prediction error effect on the P2 might similarly reflect the trial-by-trial updating of the believed contingency. However, we did not acquire contingency estimates or expectancy ratings from the participants, nor did we apply a paradigm that allowed for data-driven estimation of the learning rate. The better model fit for the high learning rate might be an indication, however, for a more conscious subjective expectancy influencing the P2 because the most recent experience is taken more strongly into account.

Limitations and Outlook

The fixed learning rates in the prediction error calculation may be considered a limitation of this study. We cannot account for interindividual differences or differences between conditions in the learning rate, and so the results of the prediction-error model comparisons should not be overinterpreted. A paradigm combining the tone-generation paradigm from sensory-attenuation studies and a forced-choice paradigm from reinforcement learning could be applied in a follow-up study to validate our results. Further manipulations regarding the prediction strength could be done, such as increasing the contingency or altering the tone identity to investigate misprediction (Hsu et al., 2018; Knolle et al., 2013b). In addition, the paradigm could be adapted to disentangle effects of attention for the uncertain action outcomes. One possibility for this is manipulating the task (ir)relevance of the tones, in general (Timm et al., 2013) or specifically for tones generated by another person, for example, by implementing a joint action setting (Loehr, 2022).

Conclusion

In this study, we explored the effect of trial-by-trial varying expectancies of the action outcome occurrence, reflected in the prediction error, on sensory attenuation, reflected in N1 and P2 amplitudes, in a paradigm with a 50% contingency between a tone and its preceding event. We acquired two groups of participants: one performing the eliciting action themselves and one observing a picture sequence of the action on screen. We compared N1 and P2 amplitudes of these action-generated tones to those of externally generated tones that were either visually cued (i.e., temporally predictable) or not. For both components, a prediction error based on a reinforcement learning model fitted the data better than the prediction error based on the statistics of occurrence in the most recent trials. Higher prediction errors led to larger N1 and P2 amplitudes, and this effect was strongest for self-generated tones. Such a potentially motor-dependent effect specific to action performance was not evident in the attenuation results, and shows that, although temporal predictability plays a role in the sensory attenuation effect, so do motor-related factors.

Acknowledgments

We would like to thank Farina Yildiz Adler, Nina Erdmann, Lea Marie Klingenberger, and Meike Vogt for their help in data acquisition.

Corresponding author: Sophie Egan, Universitätsstraße 1, 40225 Düsseldorf, Germany, or via e-mail: Sophie.Egan@hhu.de.

Data Availability Statement

Data and code are available at https://doi.org/10.17605/OSF.IO/3F6ZR.

Author Contributions

Sophie Egan: Formal analysis; Software; Writing—Original draft. Alexander Seidel: Software; Writing—Review & editing. Constanze Weber: Formal analysis; Writing—review & editing. Marta Ghio: Conceptualization; Writing—Review & editing. Christian Bellebaum: Conceptualization; Methodology; Writing—Review & editing.

Funding Information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals 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 = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). 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. The authors of this paper report its proportions of citations by gender category to be: M/M = .385; W/M = .359; M/W = .051; W/W = .205.

REFERENCES

Albrecht, C., & Bellebaum, C. (2021). Disentangling effects of expectancy, accuracy, and empathy on the processing of observed actions. *Psychophysiology*, *58*, e13883. https://doi.org/10.1111/psyp.13883, PubMed: 34196017

Bellebaum, C., Jokisch, D., Gizewski, E. R., Forsting, M., & Daum, I. (2012). The neural coding of expected and unexpected monetary performance outcomes: Dissociations between active and observational learning. *Behavioural Brain Research*, *227*, 241–251. https://doi.org/10.1016/j.bbr.2011.10.042, PubMed: 22074898

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: Event-related brain potentials in active and observational learning from feedback. *Cerebral Cortex*, *20*, 2874–2883. https://doi.org/10.1093/cercor/bhq038, PubMed: 20308202

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). Effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, *5*, 2815. https://doi.org/10.21105/joss.02815

Charpentier, C. J., & O'Doherty, J. P. (2018). The application of computational models to social neuroscience: Promises and pitfalls. *Social Neuroscience*, *13*, 637–647. https://doi.org/10.1080/17470919.2018.1518834, PubMed: 30173633

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181–204. https://doi.org/10.1017/S0140525X12000477, PubMed: 23663408

den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 548. https://doi.org/10.3389/fpsyg.2012.00548, PubMed: 23248610

Dogge, M., Custers, R., & Aarts, H. (2019). Moving forward: On the limits of motor-based forward models. *Trends in Cognitive Sciences*, *23*, 743–753. https://doi.org/10.1016/j.tics.2019.06.008, PubMed: 31371239

Egan, S., Ghio, M., & Bellebaum, C. (2023). Auditory N1 and P2 attenuation in action observation: An event-related potential study considering effects of temporal predictability and individualism. *Biological Psychology*, *180*, 108575. https://doi.org/10.1016/j.biopsycho.2023.108575, PubMed: 37156324

Eitam, B., Kennedy, P. M., & Tory Higgins, E. (2013). Motivation from control. *Experimental Brain Research*, *229*, 475–484. https://doi.org/10.1007/s00221-012-3370-7, PubMed: 23288323

Ghio, M., Egan, S., & Bellebaum, C. (2021). Similarities and differences between performers and observers in processing auditory action consequences: Evidence from simultaneous EEG acquisition. *Journal of Cognitive Neuroscience*, *33*, 683–694. https://doi.org/10.1162/jocn_a_01671, PubMed: 33378242

Ghio, M., Scharmach, K., & Bellebaum, C. (2017). ERP correlates of processing the auditory consequences of own versus observed actions. *Psychophysiology*, *55*, e13048. https://doi.org/10.1111/psyp.13048, PubMed: 29266338

Han, N., Jack, B. N., Hughes, G., & Whitford, T. J. (2022). The role of action-effect contingency on sensory attenuation in the absence of movement. *Journal of Cognitive Neuroscience*, *34*, 1488–1499. https://doi.org/10.1162/jocn_a_01867, PubMed: 35579993

Harrison, A. W., Hughes, G., Rudman, G., Christensen, B. K., & Whitford, T. J. (2023). Exploring the internal forward model: Action-effect prediction and attention in sensorimotor processing. *Cerebral Cortex*, *33*, 9130–9143. https://doi.org/10.1093/cercor/bhad189, PubMed: 37288477

Harrison, A. W., Mannion, D. J., Jack, B. N., Griffiths, O., Hughes, G., & Whitford, T. J. (2021). Sensory attenuation is modulated by the contrasting effects of predictability and control. *Neuroimage*, *237*, 118103. https://doi.org/10.1016/j.neuroimage.2021.118103, PubMed: 33957233

Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54–65. https://doi.org/10.1016/j.brainres.2015.03.038, PubMed: 25843932

Hsu, Y.-F., Hämäläinen, J. A., & Waszak, F. (2018). The processing of mispredicted and unpredicted sensory inputs interact differently with attention. *Neuropsychologia*, *111*, 85–91. https://doi.org/10.1016/j.neuropsychologia.2018.01.034, PubMed: 29407593

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, *139*, 133–151. https://doi.org/10.1037/a0028566, PubMed: 22612280

Kaiser, J., & Schütz-Bosbach, S. (2018). Sensory attenuation of self-produced signals does not rely on self-specific motor predictions. *European Journal of Neuroscience*, *47*, 1303–1310. https://doi.org/10.1111/ejn.13931, PubMed: 29738617

Klaffehn, A. L., Baess, P., Kunde, W., & Pfister, R. (2019). Sensory attenuation prevails when controlling for temporal predictability of self-and externally generated tones. *Neuropsychologia*, *132*, 107145. https://doi.org/10.1016/j.neuropsychologia.2019.107145, PubMed: 31319119

Knolle, F., Schröger, E., & Kotz, S. A. (2013a). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*, 2449–2461. https://doi.org/10.1016/j.cortex.2012.12.012, PubMed: 23318086

Knolle, F., Schröger, E., & Kotz, S. A. (2013b). Prediction errors in self-and externally-generated deviants. *Biological Psychology*, *92*, 410–416. https://doi.org/10.1016/j.biopsycho.2012.11.017, PubMed: 23246535

Knolle, F., Schwartze, M., Schröger, E., & Kotz, S. A. (2019). Auditory predictions and prediction errors in response to self-initiated vowels. *Frontiers in Neuroscience*, *13*, 1146. https://doi.org/10.3389/fnins.2019.01146, PubMed: 31708737

Kobza, S., & Bellebaum, C. (2015). Processing of action- but not stimulus-related prediction errors differs between active and observational feedback learning. *Neuropsychologia*, *66*, 75–87. https://doi.org/10.1016/j.neuropsychologia.2014.10.036, PubMed: 25446969

Korka, B., Widmann, A., Waszak, F., Darriba, Á., & Schröger, E. (2022). The auditory brain in action: Intention determines predictive processing in the auditory system—A review of current paradigms and findings. *Psychonomic Bulletin & Review*, *29*, 321–342. https://doi.org/10.3758/s13423-021-01992-z, PubMed: 34505988

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lange, K. (2011). The reduced N1 to self-generated tones: An effect of temporal predictability? *Psychophysiology*, *48*, 1088–1095. https://doi.org/10.1111/j.1469-8986.2010.01174.x, PubMed: 21261634

Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=emmeans

Loehr, J. D. (2022). The sense of agency in joint action: An integrative review. *Psychonomic Bulletin & Review*, *29*, 1089–1117. https://doi.org/10.3758/s13423-021-02051-3, PubMed: 35146702

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*, 3139. https://doi.org/10.21105/joss.03139

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*, 104092. https://doi.org/10.1016/j.jml.2020.104092

Nixon, J. S. (2020). Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition*, *197*, 104081. https://doi.org/10.1016/j.cognition.2019.104081, PubMed: 31901874

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*, 329–337. https://doi.org/10.1016/S0896-6273(03)00169-7, PubMed: 12718865

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*, 452–454. https://doi.org/10.1126/science.1094285, PubMed: 15087550

Oldfield, R. C. (1971). *Edinburgh Handedness Inventory*. APA PsycTests. https://doi.org/10.1037/t23111-000

Orpella, J., Mas-Herrero, E., Ripollés, P., Marco-Pallarés, J., & de Diego-Balaguer, R. (2021). Language statistical learning responds to reinforcement learning principles rooted in the striatum. *PLoS Biology*, *19*, e3001119. https://doi.org/10.1371/journal.pbio.3001119, PubMed: 34491980

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., et al. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y, PubMed: 30734206

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H.-J., et al. (2014). The role of reward in word learning and its implications for language acquisition. *Current Biology*, *24*, 2606–2611. https://doi.org/10.1016/j.cub.2014.09.044, PubMed: 25447993

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274. https://doi.org/10.1038/nrn2805, PubMed: 20216547

Sanmiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology*, *50*, 334–343. https://doi.org/10.1111/psyp.12024, PubMed: 23351131

Schafer, E. W. P., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science*, *181*, 175–177. https://doi.org/10.1126/science.181.4095.175, PubMed: 4711735

Seidel, A., Ghio, M., Studer, B., & Bellebaum, C. (2021). Illusion of control affects ERP amplitude reductions for auditory outcomes of self-generated actions. *Psychophysiology*, *58*, e13792. https://doi.org/10.1111/psyp.13792, PubMed: 33604896

Seidel, A., Weber, C., Ghio, M., & Bellebaum, C. (2023). My view on your actions: Dynamic changes in viewpoint-dependent auditory ERP attenuation during action observation. *Cognitive, Affective, & Behavioral Neuroscience*, *23*, 1175–1191. https://doi.org/10.3758/s13415-023-01083-7, PubMed: 36949276

Shohamy, D., Myers, C. E., Kalanithi, J., & Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience and Biobehavioral Reviews*, *32*, 219–236. https://doi.org/10.1016/j.neubiorev.2007.07.008, PubMed: 18061261

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neuroscience*, *14*, 1–11. https://doi.org/10.1186/1471-2202-14-2, PubMed: 23281832

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. https://doi.org/10.1016/j.cortex.2016.03.018, PubMed: 27137101

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. https://doi.org/10.1098/rstb.2002.1238, PubMed: 12689384

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*, R729–R732. https://doi.org/10.1016/S0960-9822(01)00432-8, PubMed: 11566114

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*, 535–544. https://doi.org/10.1093/cercor/bhh153, PubMed: 15319308

Zhang, L., Lengersdorff, L., Mikus, N., Glascher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls and suggestions of best practices. *Social Cognitive and Affective Neuroscience*, *15*, 695–707. https://doi.org/10.1093/scan/nsaa089, PubMed: 32608484

7 Appendix B: Supplementary Materials

For the published article of Study 1, a supplementary file is available online, depicting event-related potentials and pairwise comparisons. It is included in the following pages.

For the published article of Study 3, a public repository is available via the Open Science Framework (OSF). It includes files for running the experiment and files documenting the data processing and visualisation. The statistical analyses are reproducible as the data file, analysis output files, and the files for conducting the statistical analysis are also available. The link to the OSF repository is provided in the article.

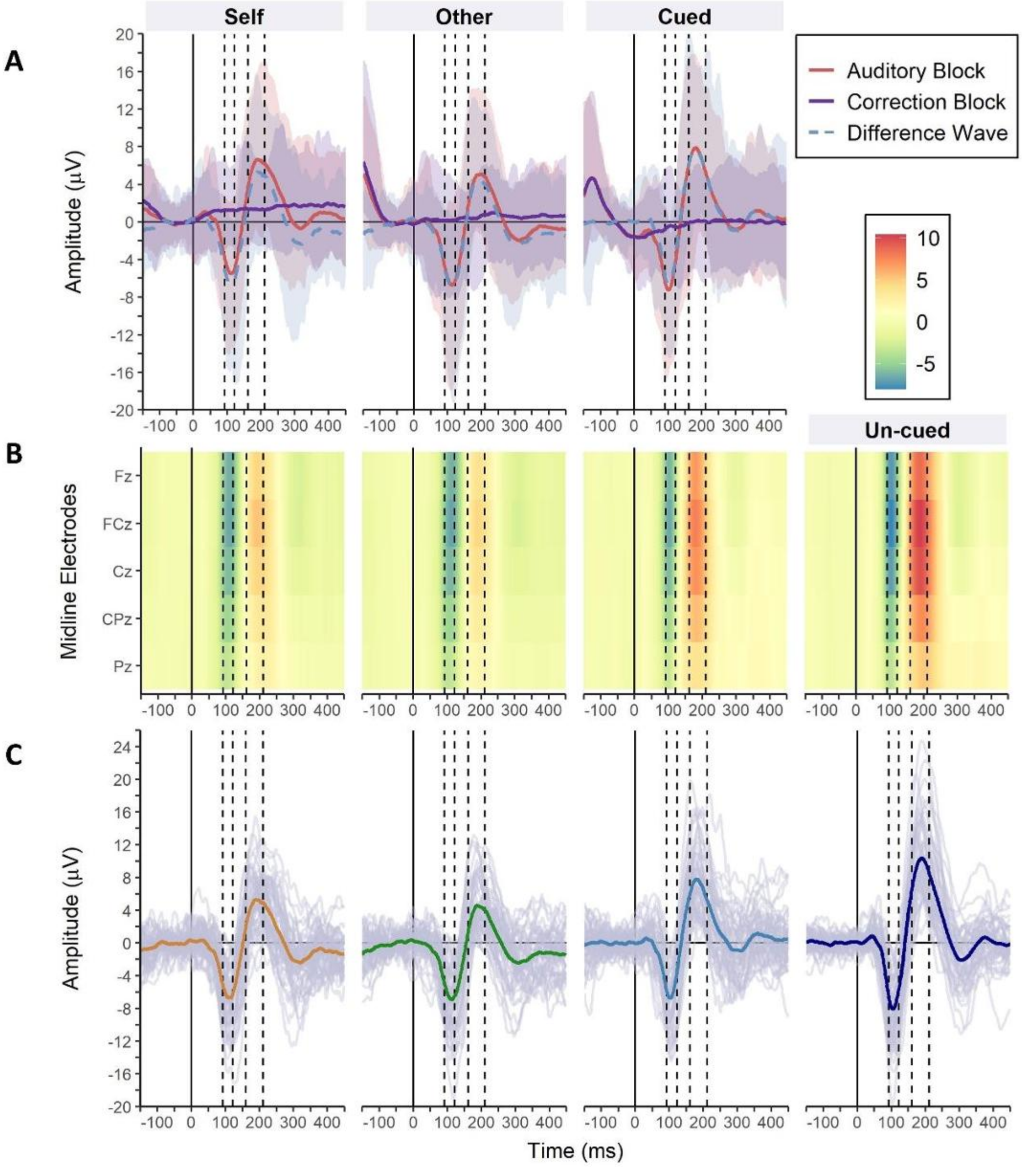
For the submitted manuscript of Study 2, an OSF repository was prepared as well. It includes more extensive documentation overall. The experimental stimuli and the files for their creation are uploaded. All aspects of the data analyses conducted in the study are reproducible, as the repository contains the anonymised raw data files, files documenting and files for conducting all steps of data processing, statistical analysis, and visualisation. This OSF repository will also be made publicly available.

Auditory N1 and P2 Attenuation in Action Observation: An Event-Related Potential Study Considering Effects of Temporal Predictability and Individualism

Sophie Egan, Marta Ghio, and Christian Bellebaum

Figure A1

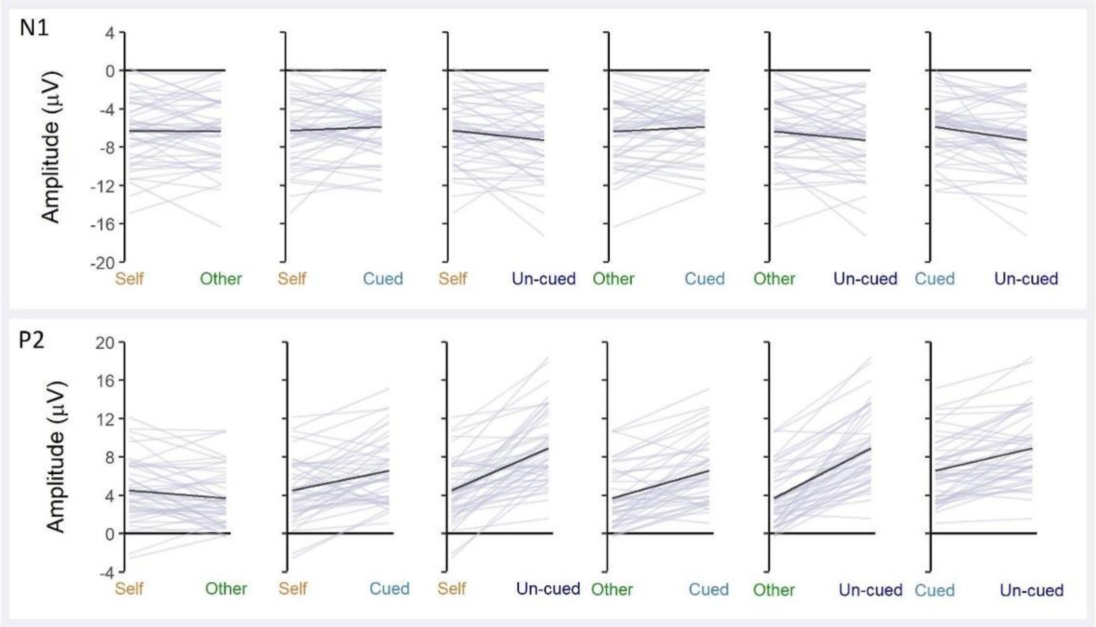
Motor and Visual Correction, ERPs of the Midline Electrodes, and Individual ERPs



Note. For all depicted event-related Potentials (ERPs) the intervals used to calculate mean N1 and P2 amplitudes are marked by dashed lines. (A) The averaged ERPs at FCz are displayed for the respective auditory blocks (auditory-motor [performed or observed] and auditory-visual), correction blocks (motor-only and visual-only), and the difference waves that are the conditions self, other and cued. Additionally, the areas between minimal and maximal participant amplitudes are coloured. (B) The averaged ERPs of the midline electrodes are presented with colour-coded amplitudes (in µV). (C) The averaged ERPs of the conditions at FCz are displayed with the individual ERPs of the participants.

Figure A2

Visualisation of the Tone Trial Type Effect



Note. Individual N1 and P2 amplitudes are connected for each pairwise comparison of the four conditions. The estimated means for each condition are displayed in black.

8 Acknowledgments

I would like to thank everyone who contributed to and supported me during the work on this dissertation, and I wish to express my gratitude especially to the following four individuals: Firstly, I very much appreciated the kindness and understanding of my first supervisor Professor Dr. Christian Bellebaum as well as his academic guidance and feedback throughout. Secondly, I would like to emphasise the impact my additional supervisor Dr. Marta Ghio had on my academic development – starting as my ExPra instructor, then as first supervisor of my bachelor and master theses, and as continuous collaborator on subsequent projects – for which I am most grateful. Thirdly, Dr. Laura Bartylla was always readily available for exchange and support, particularly regarding our common teaching experience. Fourthly, I gratefully acknowledge the constructive interactions and feedback from my office partner Constanze Weber, whose proximity, both spatial and research-related, meant she received most of my questions. More generally, I am further grateful for the exceedingly pleasant working environment my nice, cake-loving colleagues in the department of Biological Psychology provided as well as the friendly and enriching contact with colleagues from different institutes and faculties I met during the (teaching) workshops I visited. My appreciation also extends to Professor Dr. Eckart Zimmermann for agreeing to act as my second supervisor. Regarding the projects included in this dissertation, I would like to thank Manfred Mittelstaedt for creating the mirror construction used in Study 1, Jürgen Seidel and Alexander Seidel for their part in programming the experiments and scripts for data preparation, my students who contributed to the data acquisitions (Study 1: Thu Trang Vu, Study 2: Hannah Maria Kirchberg, Study 3: Farina Yildiz Adler, Nina Erdmann, Lea Marie Klingenberger, Meike Vogt), and all the participants who provided the data. Finally, I cannot overstate the importance of my family and their support and encouragement.