



UNIVERSITÀ DEGLI STUDI DI PADOVA

Department of General Psychology

Master's Course in Cognitive Neuroscience and Clinical Neuropsychology

Final Dissertation

**When expectations do not reflect reality: do event-related potential amplitudes
for self-generated sounds reflect auditory prediction errors?**

Supervisor:

Professor Umberto Castiello

Co-supervisor:

Professor Christian Bellebaum

Candidate: Valentina Mogni

Student ID Number: 2010752

Academic Year: 2021/2022

Index

Index	3
1. Introduction	5
1.1 Humans as anticipatory agents	5
1.2 Predictive processing: an overview of the main models	8
1.3 Structural and functional bases of predictive processing	13
1.4 The electrophysiological correlates of predictive processing	16
1.5 Methodology of auditory predictive processing studies	20
Abstract	27
2. Methods	29
2.1 Participants	29
2.2 Procedure	29
2.3 Material	30
2.4 Conditions	33
2.5 EEG data acquisition and analysis	36
3. Results	41
3.1 ERPs	41
3.2 Grand averages and topographies	44
4. Discussion	47
4.1 Interpretation of the results	47
4.2 Limitations	51
4.3 Conclusions	54
5. Appendix	55
5.1 Electroencephalography: an overview	55
5.2 The free energy principle	56

5.3 The Edinburgh Handedness Inventory	59
5.4 R scripts for statistical analyses	60
5.5 Mixed linear Models performed on R studio	64
References	65
6.1 Bibliography	65

Introduction

1.1 Humans as anticipatory agents

The notion that the world is our brain is embraced by several philosophers within the theories of phenomenology, giving credit to an idea of reality shaped by cerebral computations that generate a subjective and unique interpretation of it. However, perceiving was not always intended with this connotation; outdated perceptual theories portray it as a recognition process in which the senses act as a bridge from the environment to the perceiver. The so-named problem of perception has had great relief from ancient times, and many are debates related to it (Malpas, J., "Donald Davidson", 2012). The first revolutionary interpretation of the mechanisms underpinning perception arises from the *New Theory of Vision* by George Berkley (1709), which declines the act of perceiving as a passive collection of sensory stimuli. According to Berkley, motor and cognitive operations make it possible to interact with the percept and are essential for a proper understanding of the external world (Berkeley, 1709). Therefore, perceiving is an act of discovery instead of a process in which the subject is just an effortless receiver. The *New Theory of Vision* has been highly influential for the philosophers of the Seventeenth Century: following the footsteps of Berkley, the German physicist Von Helmholtz laid the foundation for the modern concept of perception. In one of his seminal writings, he postulated that: "*Each movement by which we alter the appearance should be thought of as an experiment designed to test whether we have correctly understood the invariant relations of the phenomena before us, that is their existence in definite spatial relations*" (Von Helmholtz, 1878). Highlighting the role of the subject as an experimenter, Von Helmholtz interprets perception as the product of statistical and cognitive inferences forged by experience; thus, besides acting upon sensory stimuli, humans also make predictions about upcoming situations.

The role of predictions in perception is explored by predictive processing, a theoretical framework from Computational and Cognitive Neuroscience that portray the perceptual experience as "*The use of a unified body of acquired knowledge*" (Clark, 2015). A dominant view from predictive processing theories is that predictions arise from the continuous interactions between bottom-up signals from the environment and top-down signals, working as anticipated inputs. Predictive processing allows distinguishing self-produced stimuli from external stimuli across different sensory modalities; indeed, the cortical response to self-generated stimuli is less intense than the response elicited by the same stimuli when generated independently from one's behavior (Schafer & Marcus, 1973; Blakemore et al., 1999; Lindner

et al., 2005; Sato, 2008). The reduction of cognitive resources directed toward self-generated stimuli is adaptive: because they come from actions that the subject performs voluntarily, there is no need to process their consequences, and the saved resources can be addressed for other tasks or for processing other stimuli coming from the outside, whose effects are less predictable and may require an immediate response (Dogge et al., 2019). Although predictions allow us to achieve multiple goals, most studies focus on their role within the visual system, which takes advantage of them for several purposes, such as maintaining the stability of the output across eye movements and creating a coherent percept (Sommer & Wurtz, 2008). In the auditory system, predictive processing enables the recognition of self-generated sounds, with particular attention to speech (Wang et al., 2014) and to sounds triggered by simple actions of one own (SanMiguel et al., 2013). Accordingly, predictions of somatosensory stimuli make it possible to recognize and differently process self-generated tactile sensations, leading to a diminished activation of somatosensory areas for their high predictability; this is also why it is so difficult to tickle ourselves (Bays et al., 2006; Blakemore et al., 1999).

However, as the word itself suggests, predictions are just predictions and can be disproven. The concept of Prediction Error (PE) arises from this assumption: theoretically, PEs reflect the mismatch between expectations and reality (Friston, 2005). Predictions are experience-based, and they rely on the extraction of statistical regularities from everyday life situations used to create an internal and coherent representation of the world: when expectations mirror reality, prediction errors are minimized; when expectations deviate from it, they are consistent and signaled by specific cortical activity, indicating the need to update the internal model of the world (Den Ouden et al., 2012). Furthermore, PEs are heterogeneous and serve different functions, being part of a general neural coding mechanism aimed to build an internal model of the world. They have a role in shaping the perceptual experience, but they are also a key concept when referring to reward, motivational control, and decision-making (Den Ouden et al., 2012). The ability to constantly update world representations allows for development and adaptation; conversely, some neuropsychiatric disorders manifest severe impairments in making predictions that derive from prediction error signaling dysfunctions. Among the dimensional categories of DSM-V, two psychiatric disorders characterized by impairments in Predictive processing are Obsessive-Compulsive Disorder (OCD) and Schizophrenia (*DSM-5 Manuale diagnostico e statistico dei disturbi mentali - American Psychiatric Association - Raffaello Cortina Editore - Libro Raffaello Cortina Editore*). When referring to OCD, failures of Predictive processing result in a loss of the ability to rationalize the consequences of self-made

actions, which are usually overestimated; this translates into notable distortions of the sense of agency directly linked to the presence of compulsions and obsessions (Gentsch et al., 2012). A consistent number of studies have also investigated how Predictive processing failures can result in auditory hallucinations and cognitive delusions, namely two of the hallmark symptoms of Schizophrenia (Bansal et al., 2018). More in-depth, the inability to properly weigh prediction errors can result in a psychotic episode when the subjective inferences are no more distinguishable from the objective experience of the moment (Ford et al., 2014).

Overall, predictive processing explains how each person can shape the world firsthand. It highlights the role of a priori computation in guiding perceptual processes; the subject is an active perceiver who must learn efficient predictive mechanisms to function in the environment. In the following paragraphs, we review the main theoretical models that reveal the mechanisms of predictive processing and the role of prediction errors. The electrophysiological correlates of prediction errors in the auditory system are examined in the current study; for this reason, the introduction focuses on the cognitive computations that define predictions in this specific domain so that the reader can get a better insight into the theory of the subject. In addition, paragraph 1.3 explains the anatomical connections between the motor and the auditory cortex, the brain regions activated during the processing of motor-cued auditory stimuli. Focusing on auditory predictions is convenient because they are easier to study than predictions in other sensory modalities: there is clear evidence for the existence of a pathway connecting motor and auditory brain regions, but the same evidence does not exist for the visual domain. Moreover, experimental paradigms studying predictions of auditory stimuli are easier to create for the possibility of distinguishing predictions from the elaboration of environmental stimuli; a clear distinction between the two can only be inferred in other domains, such as the somatosensory domain (Reznik & Mukamel, 2019).

1.2 Predictive processing: an overview of the main models

Predictive processing is related to a broad range of computations, and a model that depicts its precise functioning is hard to accomplish. Classic theories of perception rely on hierarchical models, assuming that stimuli undergo different processing steps that extract increasingly complex features (Riesenhuber & Poggio, 1999). Hierarchical models are consistent with the anatomy and physiology of the cortex, whose structural organization in layers reflects the different processing steps of the hierarchy. With Predictive processing, hierarchical models transform into the so-called hierarchical generative models; the latter differs from the previous ones for the presence of *a priori* cortical signals embodied in the hierarchy, representing predictions. Generative models are related to the cognitivist concept of ‘analysis by synthesis’ (Neisser, 1967), which explains perception as a process of adaptation of internal models to match the actual sensory input. The influence of top-down expectations on bottom-up signals is the key, and it is possible thanks to the existence of recurrent connections between the cortical layers (Mumford, 1992).

Internal forward models have been the reference ones to explain predictive processing for a long time due to their emphasis on recurrent connections. They assume that the neural network activated during perceptual processing contains different units and that Efference Copies (ECs) are the main ones, carrying the necessary information to predict upcoming events. Efference copies are the starting point for making inferences about self-cued sensory stimuli since they contain a copy of the stimuli resulting from motor commands used by the subject to explore the environment (Bansal et al., 2018). Therefore, the association between the upcoming stimulus and the action can be used during perceptual processing to infer sensory predicted consequences based on experience. This information comes from higher processing levels, going down the processing hierarchy where it encounters the upcoming sensory signals at lower levels. Internal forward models use the adjective “high” to indicate processing levels that do not directly process information from the environment, making inferences about it instead. The product of the computations is a predicted sensory consequence, available even before the initiation of a movement, which depends on motor intention rather than motor execution (Crapse Sommer, 2008; Reznik et al., 2018). Then, the prediction is compared with the actual sensory stimulation, and the predicted hypotheses that do not match reality are sent back to higher processing levels; feedforward connections are fundamental in sending the discrepancy to the upper layers to adapt the motor command to the new expectation. The internal network keeps updating until it reaches a convergence, producing a consistent embodiment of reality: internal forward models

intend convergence as the point to which Efference Copies overlaps upcoming stimuli (Walsh et al., 2020). Due to their influential role in this iterative process, feedforward connections are the key to perceptual adaptation and skills acquisition (Dogge et al., 2019). Moreover, the motor information stored by Efference Copies allows for a distinction between stimuli coming because of self-made actions and stimuli caused by external factors, shaping the personal notion of self-agency (Gentsch & Schütz-Bosbach, 2011).

Classic forward models rely exclusively on Efference Copies to explain predictive processing. According to these models, predictive processing arises from the continuous comparison between the inferred sensory consequences expected from a given action and the actual upcoming stimuli; the constant update of the information stored in Efference Copies allows the subject to be more accurate in future explorations. Other theoretical accounts, such as Predictive Coding (PC) or Ideomotor theories, view predictive processing as a more general predictive mechanism in which perceptual models do not stick to motor commands only but arise from a prediction error minimization process. Therefore, every kind of active inference made by the subjects has a role in perception (Dogge et al., 2019a). This difference is crucial because it implies the presence of predictions and prediction errors in models of reward, learning, motivational control, attention, and orientational skills (Den Ouden et al., 2012). By proving the presence of the electrophysiological correlates of predictive processing without motor cues, Kaiser et al. (2018) bring evidence for more general prediction mechanisms. In particular, the authors underline the role of attentional salience and motivational control in shaping the predictive response (Kaiser & Schütz-Bosbach, 2018). Other variables, such as task rules, have been proven to shape predictions bringing additional evidence for a general predictive mechanism (Dogge et al., 2019a). These findings lie in the notion that PE should be a common currency across brain regions, but the nature of the signal changes according to its precise role (Dogge et al., 2019b).

Rao and Ballard (1999) formalized the Predictive Coding (PC) theory to explain action-unrelated predictions, and they have accomplished this goal by introducing the concept of prediction errors. Building on their studies on the visual system, the authors have compared the anatomical organization of the visual cortex to the structure of a hypothetical predictive processing network, organized in layers containing forward, backward and lateral connections (Maunsell & Essen, 1983). The forward ones run from lower to higher layers, while the backward ones run from higher to lower layers; lateral connections connect units within the same laminar level, allowing for inhibition of the signals that do not need to be sent backward during PP (Friston, 2005). Starting from these considerations, Rao and Ballard hypothesize a

model in which each hierarchical level includes two types of mutually-connected units: Expectation (Ex) and Error (Er) units. Ex units carry predictions about the upcoming stimuli and send them to Er units. As in every generative model, if the hypotheses carried by Ex do not match reality, prediction errors are feedforward from Er to the Ex units of the preceding layers, communicating the need to revise expectations to minimize prediction errors (Rao & Ballard, 1999). As abovementioned, lateral connections are in charge of inhibiting the overlapping information in Er and Ex. Indeed, the units carrying the same information produce a null PE. Since a revision by Ex is unnecessary when PE equals zero, the signal does not need to be fed forward to upper levels. The schematic representation of the model in Figure 1.1 allows us to observe how neural populations are mutually connected both inside the same level of the hierarchy and outside, communicating with the other layers.

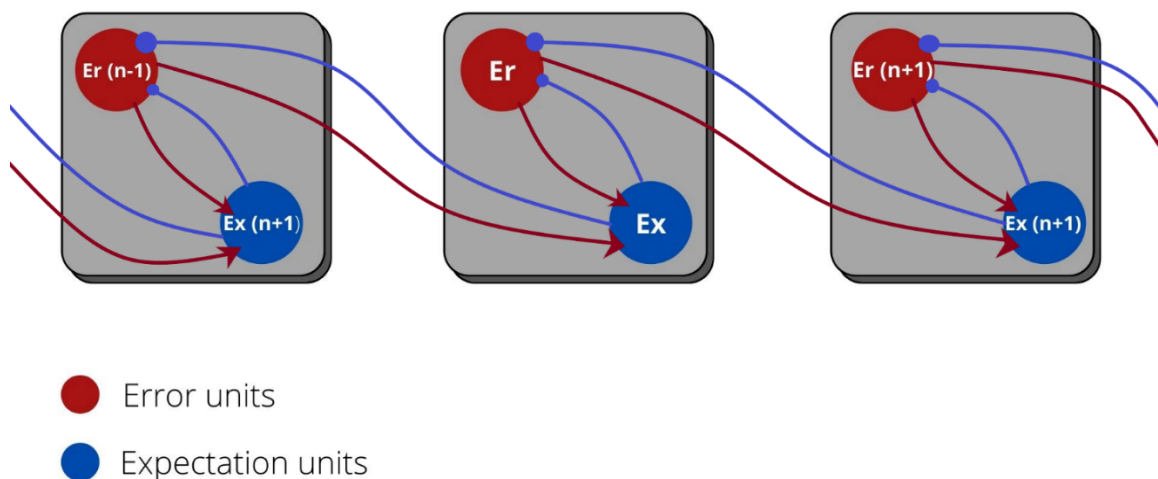


Figure 1.1 Schematic illustration of the model by Rao and Ballard (1999). The grey squares represent the layers of the hierarchical network, each containing Expectation Units (Ex) and Error Units (Er). Ex units communicate with Er units of the preceding layer, sending predictions via backward connections. Er units send prediction error signals, reaching the previous Ex units via feedforward connections. Moreover, the units inside each layer are in mutual communication, allowing for the internal exchange of information (Rao & Ballard, 1999).

Whereas Rao and Ballard have conceptualized the hypothetical structure of the predictive coding network, Friston (2009) has tried to develop a mathematical formulation for its computations. The author has built his theory on the Bayesian brain hypothesis, based on the principle

that the brain is comparable to a computational machine (Knill & Pouget, 2004; Friston, 2009). Therefore, predictive processing can be seen as an empirical Bayes mechanism since it starts from *a priori* factors to produce the final output. The formula reported in section 5.2 of the Appendix is an attempt to parametrize the computations occurring within a predictive processing network: every prediction assumes a probabilistic value, represented by the product between the previous experience of the event (Vp) and a given variability index (Sp) (Friston, 2005). Such probability values are constantly updated to fit the upcoming stimuli; thus, the output value given by the formula should represent a realistic picture of reality, considering all the plausible hypotheses with their respective probabilities. In addition, Friston states that the final goal of predictive coding is to minimize the amount of free energy inside the system (Friston, 2009). In this context, free energy is the output of a function that links sensory data and brain states; it is related to the amount of entropy, which in turn represents the amount of surprise experienced by the subject during the joint occurrence of a sensory event and its cause. Thus, the energy reduces when the surprise level of the perceiver is low, meaning that the system has learned how to explain a specific situation (Friston, 2009).

Last, Ideomotor theories take another perspective to explain predictive processing. They start from the assumption that motor planning and the processing of motor outcomes rely on the same functional bases. Therefore, motor commands and their sensory effects are stored in the same units and can reciprocally influence each other. Nonetheless, Ideomotor theories have mainly explained the influence of external stimuli on action planning and are seldom used to account for the inverse relationship. However, this does not preclude using them to explain how actions affect perception. Indeed, they constitute one of the main theoretical accounts of Motion-Induced Blindness (MIB): the effect implies that a moving pattern is voluntarily neglected by the subject when it does not bound to a previously learned motor pattern (Bonneh et al., 2010). The effect is very similar to sensory attenuation; thus, this bodes well for the possibility of using such theories to explain predictive processing.

However, since predictive elaborations are widespread inside the brain, the theories need to be generalized to other domains besides the motor. Indeed, prediction errors are not only involved in explaining action-related outcomes, but their role has been detected in cognitive, learning, motivational, and attentional processes (Den Ouden et al., 2012; Bansal et al., 2018). Even though internal forward models have been a good starting point for the study of predictive processing, they present some theoretical constraints first raised in animal studies. For instance, the modification of the neural pathways related to novel actions learned to explore the

environment is relatively slow (Schneider et al., 2018); in humans, this modulation occurs in a fast way even when the person has not had a solid experience about the context (Dogge et al., 2019). Therefore, there could be more than one mechanism behind predictive processing, modulating different neural populations at different velocities. Dogge and colleagues (2019) suggest that hybrid prediction models can be an optimal theoretical account: the models provide alternative pathways driving predictions. The path choice depends on the objective, is based on previous experience, and relies on multiple cognitive resources, such as beliefs, memories, and intentions (Dogge et al., 2019). Recent studies have highlighted the relevance of precision-weighted predictive signals, with precision being a factor that can upgrade predictive processing networks. Functionally speaking, the level of precision within the model derives from the accuracy of the lateral inhibitions occurring during feedback and feedforward signaling (Friston, 2018).

Moreover, the state of the art in predictive processing views it as a holistic mechanism that encompasses learning, attention, memory, and sensory attenuation (Friston, 2018), with prediction error minimization mechanism remaining at the core of it. The following paragraphs delve into the structural and functional bases of predictive processing, focusing on prediction error minimization in the auditory domain. However, it is essential to consider that predictive processing occurs in conjunction with other processing, allowing for a coherent and complete experience of the world.

1.3 Structural and functional bases of predictive processing

Predictive processing reflects the structural organization of the cortex, which consists of laminar layers (Barbas & Rempel-Clower, 1997) and feedback and feedforward connections (Lamme & Roelfsema, 2000). Additionally, the functional principles of predictive processing mirror the general functioning of the sensory cortex, and functional Magnetic Resonance Imaging (fMRI) studies have shown that different phases of predictive processing rely on separate laminar profiles (Heilbron & Chait, 2018). On the strength of internal forward models, action-based predictions arise in the anterior cerebellar cortex and rely on the signal from supplementary motor areas (Haggard & Whitford, 2004). When the cerebellum receives the motor information and creates predictions, those latter reach the parietal sensory area of interest through different corticothalamic pathways (Bansal et al., 2018). In addition, the weight of the prediction error signals that define the difference between predicted and actual stimuli depends on tonic dopaminergic firing within these pathways (Friston et al., 2012). Furthermore, it is relevant to consider the numerous connections between motor and sensory areas across modalities that characterize different predictions; however, the most studied cerebral connections are the thalamocortical and cerebellar ones (Mumford, 1992; Parras et al., 2017; Woodward et al., 2012). The thalamocortical networks are commonly organized in parallel, connecting different cortical regions to specific thalamic nuclei (Alexander et al., 1986); feedback and feedforward connections allow communication in both directions. Furthermore, failures in prediction error signaling relate to dysfunctions that can appear at every level, affecting the generation, the transmission, or the utilization of the signals coming from Efference Copies (Bansal et al., 2018). Signaling failures within the thalamocortical-cerebellar pathway that are persistent over time manifest in the typical symptoms of schizophrenia, defined as Cognitive Dysmetria, because of the inability to organize information and create coherent representations of the perceptual world (Andreasen, 1997). The neuropathology of the disorder has been associated with altered functional connectivity between the thalamus and the cortex: schizophrenic patients manifest a decreased resting-state connectivity between prefrontal and thalamic areas, simultaneously with an increased resting-state connectivity between thalamic nuclei and prefrontal cortices (Woodward et al., 2012). If the affected individuals retain some reality-checking capacities, disruptions in thalamocortical signaling manifest with milder symptoms. For instance, they may result in a synesthetic experience when cerebellar signals reach more than one sensory cortex, such that multi-modal predictions are

activated and manifest simultaneously, enriching the perceptual experience with unusual sensations (Seth, 2014).

As mentioned before, there is evidence of a motor-induced sensory modulation for all sensory modalities (Bansal et al., 2018; Blakemore et al., 1999; Schütz-Bosbach & Prinz, 2007) even though, up to date, there is no proof of a clear anatomical connection between primary motor and sensory areas. Functional Magnetic Resonance Imaging (fMRI) studies evidenced a raising in the activity of the Heschl's gyrus (A1), the superior temporal gyrus (STG), the inferior temporal gyrus (IFG), the motor regions M1 and SMA, and the cerebellum during the elaboration of motor-cued auditory stimuli (Haggard & Whitford, 2004; Reznik et al., 2015a). Furthermore, studies using this neuroimaging technique have evidenced a functional coupling between the motor and the auditory system, suggesting the presence of subcortical connections between the areas, most probably connecting the thalamus, the brainstem, the cerebellum, and the cortices (Reznik et al., 2014; Reznik et al., 2015a). Concerning the cortical pathways, animal studies point out the presence of bi-directional connections between secondary motor and auditory areas but no connections between the primary ones (Parras et al., 2017). In human studies, most findings arise from Diffusion Tensor Imaging (DTI), although this technique does not allow for investigating the directionality of connections. DTI proves a direct linkage between the frontal cortices and the STG, connected via the arcuate and superior longitudinal fasciculus; moreover, the motor IFG and the auditory STG connect through the extreme capsule (Frey et al., 2008; Saur et al., 2008); figures 1.2(a) and 1.2(b) illustrate the main white matter tracts connecting the brain regions within a plausible predictive processing network, evidenced by diffusion tensor tractography.

The first units containing error-related signals appear within the Inferior Colliculus (IC) (Parras et al., 2017), while the secondary auditory cortex (A2) is where prediction errors manifest their highest expression (Carbajal & Malmierca, 2018). Therefore, subcortical pathways have an essential role in predictive processing. The induced deactivation of cortical structures, achieved with neurostimulation techniques, can highlight the relevant subcortical dynamics (Carbajal & Malmierca, 2018). Also, it is essential to consider the presence of bidirectional connections between the regions, with different neurotransmitters modulating the transmission of the signal. The GABA neurotransmitters have a prominent role since they regulate the inhibitory activity of feedback and feedforward connections (Ayala & Malmierca, 2015). Because of their rapid ionotropic action, GABA-A receptors are most likely to directly influence predictive processing, while GABA-B receptors should be only secondarily involved, affecting predictive

processing with a slow metabotropic effect (Duque et al., 2014). A cholinergic modulation, arising from the Inferior Colliculus and controlled by muscarinic receptors, allows the allocation of attentional resources towards the stimuli to elaborate (Ayala & Malmierca, 2015). It has a fundamental role in updating the internal model, allowing for a precise selection of the features to revise by allocating more resources towards them (Abolafia, Vergara, Arnold, Reig, & Sanchez-Vives, 2011; Sanchez-Vives, Nowak, & McCormick, 2000a, 2000b).

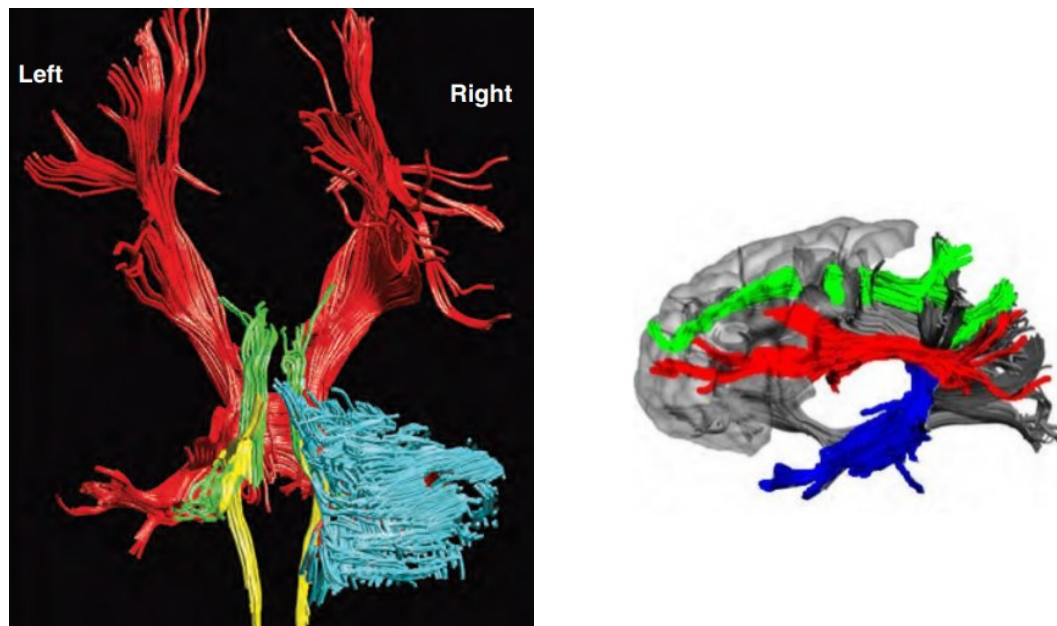


Figure 1.2 (a) and 1.2 (b) Main white matter tracts involved in the network. Figure 1.2 (a) (left) shows the main cerebellar connections. The green tract represents the superior cerebellar peduncle, connecting the anterior cerebellar cortex to the thalamus and the cerebral cortex. Figure 1.2 (b) (right) highlights the lateral part of the arcuate fasciculus (blue) and two main tracts of the superior longitudinal fasciculus (red and green). The figures are taken and modified from the Atlas of Human Brain Connections.

Catani, Marco, and Michel Thiebaut de Schotten, *Atlas of Human Brain Connections* (Oxford, 2012; Oxford Academic), <https://doi.org/10.1093/med/9780199541164.001.0001>

1.4 The electrophysiological correlates of predictive processing

Predictions act as a cognitive filter for relevant information and enable the creation of a coherent representation of the world. Several studies specifically address the ability of the predictive processing network to store regularities between environmental stimuli (Bendixen et al., 2012). For example, voluntary actions can be associated with specific sounds if they frequently occur simultaneously. Their processing differs from the processing of external sounds: self-cued auditory stimuli are subject to the “self-stimulation” or “sensory attenuation” effect (Schafer & Marcus, 1973), for which the brain response to a self-generated sound is attenuated compared to the response elicited by the same sound when generated by an external source (Schafer & Marcus, 1973). The self-stimulation effect is more pronounced in the auditory rather than the visual domain, and the attenuation of the brain activity correlates negatively with the intelligence scores of the subject as measured with the IQ test scale (Schafer, 1982). The processing of self-generated sounds specifically relates to a reduction in the amplitude of some Event-Related Potentials (ERPs) (Horváth, 2015), namely the electrical potentials produced in response to specific sensory, cognitive, or motor stimuli (Luck, 2014). Schafer and Marcus (1973) report that almost all ERPs undergo attenuation in response to a self-induced stimulus (Schafer & Marcus, 1973). However, more recent studies highlight the attenuation of N1, P2, and P3 and their respective subcomponents (Horváth, 2015) as prominent. N1 and P2 are functionally independent and reflect different cognitive processes: N1 suppression links to cerebellar motor-based predictions and is absent in patients with cerebellar lesions (Knolle et al., 2013). Also, the N1 suppression effect reduces when the predictability of the sound is less clear (Bäbß et al., 2008) or the delay between the action and the sound is long (Davis & Zerlin, 1966). Accounting for feedforward models, a reduced suppression for unpredictable stimuli might link to the lack of a link between the motor and sensory information within the Efference Copies. The timing of stimuli is crucial for their attribution as self-generated or external (Schafer & Marcus, 1973). Moreover, N1 suppression is only present in concomitance to motor cues (Lange, 2011; Sowman et al., 2012); this is not the case for the P2 component, which results independent of the cerebellar activity (Knolle et al., 2013) and reflects more general elaborations such as the stimulus predictability (Neszmélyi & Horváth, 2017; Knolle et al., 2013) and the judgments about agency, namely the specific attribution of the action to the self (Synofzik et al., 2008). Therefore, N1 suppression reflects the forward model accounts and provides evidence for the role of motor-based inferences in predictive processing. On the other hand, P2 changes underlie broad predictive mechanisms that are not exclusively tied to motor

commands (Dogge et al., 2019; Friston, 2005; Rao & Ballard, 1999). Additionally, P2 modulation depends on the context in which the stimulus appears: Miniati et al. (2010) found a reduction in P2 amplitude during the processing of sounds inserted into a melody compared to tones presented alone (Miniati et al., 2010). In other words, N1 and P2 amplitude decrease when the sound that occurs after the action satisfies the predictions (Baess et al., 2008; Knolle et al., 2012). On the other hand, N1 and P2 amplitudes increase when the expectations about upcoming stimuli are not met, reflecting the signaling of prediction errors (Knolle et al., 2013). Error signaling relies on specific neural correlates; several studies demonstrate that the suppression effect is reduced when predictions are unmet (Baess et al., 2008; Martikainen et al., 2005), with deviant stimuli triggering an enhanced N1 response (Näätänen, 1990). Specific findings on these ERP components during the processing of auditory stimuli highlight an increased amplitude after unexpected feedback, ruling out the interpretation of the ERPs modulation as a simple reaction to action-stimulus contiguity (Behroozmand and Larson, 2011; Heinks-Maldonado et al., 2005). Therefore, the reduced suppression of the N1 and the P2 components correlates with the processing of prediction errors; according to internal forward models, increased auditory responsiveness to deviants improves the processing of unexpected features and increases their salience. Knolle et al. (2013) also reported changes in the N2b and P3a amplitude when deviant stimuli occur after self-generated cues and are more related to the categorization of unexpected stimuli (Knolle et al., 2013). Other electrophysiological changes correlate with the elaboration of unexpected stimuli during predictive processing to provide a label to the system when identifying the information to be updated. This is essential to revisit the cognitive assumptions inside the internal cognitive model for predictions and prevent hallucinatory experiences (SanMiguel et al., 2013). During the Oddball paradigm, used to study omission responses (see the following paragraph), a specific electrophysiological response is temporally coupled to the omitted stimulus (Raij et al., 1997). In the auditory domain, omissions signals, which represent deviant stimuli, prominently correlate with electrophysiological changes at the level of the Superior Temporal Gyrus and the secondary Auditory Cortex (Heilbron & Chait, 2018). Studies using the Oddball paradigm to investigate prediction error correlates highlight first the activation of the template for the sensory stimulus (SanMiguel et al., 2013), followed by the presentation of the Mismatch Negativity indexing the deviance (Paavilainen, 2013). Therefore, MNN was first associated with the appearance of incongruent patterns without considering its role in predictive processing. Thanks to evidence from Repetition Suppression studies (Henson, 2003), it has been possible to disentangle the

electrophysiological pattern correlating with repetition from that associated with deviant stimuli in the Oddball paradigm (Den Ouden et al., 2012). It turned out that the MNN is not related to the elaboration of the deviant stimulus per se but violated predictions. Nevertheless, MMN responses are temporally constrained and seldom present when the time interval between the sounds is under 200 ms (Rajj et al., 1997; Todorovic et al., 2011); for this reason, its functional role is still under debate. Chennu et al. (2013) attempt to find a defined electrophysiological pattern correlating with the processing of prediction errors and identify three elaboration steps: the MNN is the first component detected and correlates with the magnitude of prediction errors; 100 ms later, the P3 correlates with the error magnitude and reveals higher elaborations; additionally, the P300 changes according to the attentional resources implied in PP and depends on the strength of top-down expectations, linked to reward evaluations (Wu & Zhou, 2009). The last component reflecting prediction error processing is the Contingent Negative Variation (CNV), which carries the elaboration of broad predictions that do not refer to a specific stimulus (Chennu et al., 2013). In addition, Todorovic and Lange (2012) dissected the electrophysiological pattern underlying stimulus expectancy and repetition, finding differences in the response timing. The brain activity in the time window of 40-60 ms after the stimulus is associated with stimulus repetition responses, while later elaborations (occurring 100 ms after the stimulus) result in predictive processing (Todorovic & Lange, 2012) and thus constitute the ones of interest in the current study.

Time-frequency analysis is another valid method to capture cognitive dynamics and has been used in recent studies about auditory predictive processing (Buschman & Miller, 2007; Chao et al., 2018; Heilbron & Chait, 2018; van Kerkoerle et al., 2014). It accounts for changes in the oscillatory rhythms of different frequency bands that occur during specific cognitive processes. Typically, changes in rhythmic activity correlate with task demands, including perceptual, cognitive, motor, language, mnemonic, and other elaborations (Cohen, 2014). Feedback and feed-forward signals linked to predictive processing have been associated with distinct oscillatory signatures: the first, which transmit predictions downwards from higher layers, are mainly transmitted over beta frequencies (12-30 Hz) or lower; the second, which transmit error signals upward, propagate over the gamma band (> 30 Hz) (van Kerkoerle et al., 2014; Buschman and Miller, 2007). As with ERPs, prediction errors are also related to specific oscillatory rhythms. Cao et al. (2017) examine brain oscillations during the predictive processing of self-generated sounds. The study replicates previous findings on the suppression effect and provides evidence for a general decrease in electrophysiological activity during the processing of self-cued

auditory stimuli. Furthermore, the authors have identified some pre-stimulus changes within the alpha band (8-12 Hz) that affect the elaboration of the sounds and may be responsible for the sensory attenuation effect, carrying expectations about upcoming events. They also pointed to a correlation between alpha/beta and gamma time-frequency changes, which reflects different steps of predictive processing and the exchange of signals between feedback and feedforward connections. In addition, a phase-locking analysis evidenced some differences between the passive and the active condition, namely the condition in which the subject actively cues the sound or processes it as an external stimulus. When the expectations are unfulfilled, a significant reduction in alpha and beta phase-locking activity has been registered for the first condition but not for the second one (Cao & Gross, 2015). Additionally, the idea of alpha and beta rhythms acting as top-down signals has received support from other studies: Frontolan et al. (2014) and Arnal et al. (2011) demonstrate that the alpha and beta oscillations modulate the gamma power activity in the primary auditory cortex; this can reflect the transmission of prediction error signals to solve prediction errors (Arnal et al., 2011; Fontolan et al., 2014). The current study will investigate N1 and P2 changes during predictive processing when different types of cues predict the sound. The electrophysiological signatures abovementioned are not considered as the variables of interest, although a brief overview of them helps to get an overall picture of the understanding of predictive processing, and it helps to raise awareness that N1 and P2 are not the only electrophysiological components to change during predictive processing; in particular, the P3a could have been studied as paired with the P2, having a functional correlation with the memory and attentional processes involved in predictive processing.

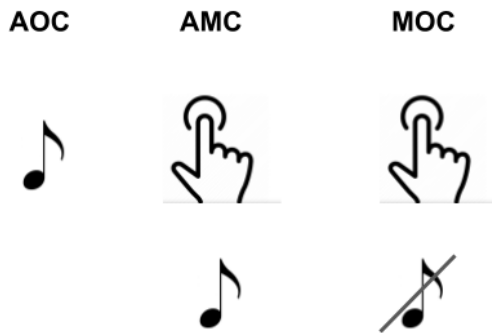
1.5 Methodology of auditory predictive processing studies

Initial paradigms for the study of predictions. The Oddball paradigm is the first and most robust experimental paradigm for studying the electrophysiological correlates of deviant stimuli (Squires et al., 1975). It examines visual, auditory, cognitive, and emotional computations (Halgren et al., 1998; Schlüter & Bermeitinger, 2017) during repeated presentations of the same stimulus, occasionally interrupted by a deviant one. The previous paragraph describes the importance of the Oddball paradigm in demonstrating the occurrence of the Mismatch Negativity (MMN) in processing deviant stimuli. Studies of repetition suppression have been fundamental in disentangling the electrophysiological pattern for incongruent stimuli from that of predictive processing of the occurring stimuli. The Repetition Suppression paradigm can take various forms and consists of the repeated presentation of a stimulus, which can be either cued or not. Based on the findings from these two paradigms, a broad branch of studies focusing on the electrophysiological correlates of predictive processes has differentiated the computations in each sensory modality. Predictions related to the auditory domain are mainly studied in relation to speech sounds to distinguish self-generated from external sounds (Numminen et al., 1999; Näätänen et al., 2010) or sounds appearing after simple actions, such as a button press (Horváth, 2015; SanMiguel et al., 2013). As described in the methods, the current study focuses on the second type of stimuli, analyzing the electrophysiological correlates of prediction errors during processing sounds that may appear after a button press.

Training sessions. Most studies investigating prediction errors include training sessions prior to the experimental conditions, strengthening the association between sounds and cues in the internal cognitive model. Training helps to learn the rhythm of button presses, as the timing of stimuli is essential for the correct detection and registration of ERP signals (Horváth, 2015; Luck, 2014); it also helps to distinguish between body- and environment-related predictions: For the latter, training is crucial to establish an association. Moreover, simultaneous registration of the EEG signal shows ERP suppression when self-generated sounds occur (Baess et al., 2011; Sanmiguel et al., 2013; Schafer & Marcus, 1973; Timm et al., 2016; review Horváth, 2015). Studies using magnetoencephalography support EEG studies and highlight suppression of the N1 magnetic counterpart: M100 (Aliu et al., 2009; Martikainen et al., 2005).

Paradigms for the study of auditory predictive processing. The Contingent Paradigm is one of the best-known paradigms for disentangling the neural correlates of auditory predictive processing (Horváth, 2015). The paradigm assumes that self and externally generated sounds are processed differently, as demonstrated by studies investigating the sensory-suppression effect (Knolle et al., 2013). It is divided into three main conditions (Miall & Wolpert, 1996), usually administered in separate blocks: the Auditory Motor Condition (AMC), the Auditory Only Condition (AOC), and the Motor Only Condition (MOC). During the Auditory Motor Condition (AMC), the participant performs an action, such as a button press, always followed by an auditory stimulus; in the Auditory Only condition (AOC), the participant only listens to externally generated sounds either from another person (Cao & Gross, 2015; Ghio et al., 2018; Sato, 2008; Weiss et al., 2011a, 2011b) or from electronic devices (Aliu et al., 2009; SanMiguel et al., 2013; Schafer & Marcus, 1973) In the Motor Only Condition (MOC), the participant performs the button press but no sound follows. After data collection, the EEG signal recorded during the MOC is subtracted from that recorded during the AMC to obtain the motor-corrected AMC signal. The motor-corrected signal binds to inferences about the auditory stimuli appearing after the button press and gets compared to the AOC signal, in which predictions about it are not present since the sound never appears (Horváth, 2015). In short, predictive computations related to the sound are present in the AMC but not in the AOC. Although the Contingent Paradigm relates predictive processing to its electrophysiological correlates, it does not allow the specific correlates that characterize the signaling of prediction errors. To accomplish this goal, San Miguel et al. (2013) used another paradigm with similar conditions, which they call the N1 suppression paradigm, taking as reference the methods by Arnald and Giraud (Arnal & Giraud, 2012). The authors studied what they call “hearing silence” (SanMiguel et al., 2013), analyzing how the electrical correlates of predictive processing change when the sound does not always appear after the action and the subject “hears the silence” instead. Therefore, they study the neural correlates of prediction errors because top-down predictions are not fulfilled. As illustrated in Figure 1.3, the AMC and the AOC are identical to those of the Contingent Paradigm, but the MOC is absent in this case. Instead of the MOC, a condition in which the sound does not always follow the button press is present: when the prediction is available but the output does not follow, the electrophysiological signal carries the prediction error response (SanMiguel et al., 2013).

Contingent Paradigm (Horváth, 2015)



SanMiguel et al. (2013)

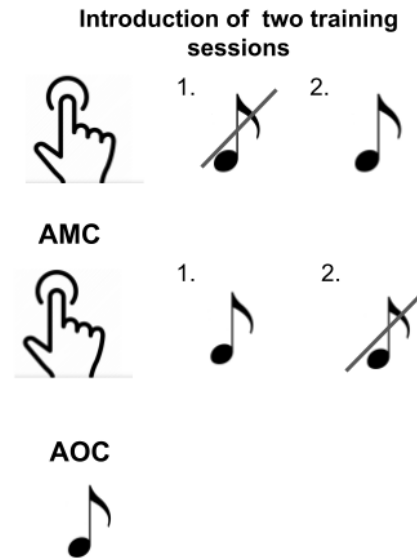


Figure 1.3 Horváth (2015) and SanMiguel et al. (2013) experimental paradigms. The second paradigm includes two training sessions: the first one (1.) requires the subject to press the button, but no sound occurs; the second training (2.) requires the subject to press the button, and the sound always occurs. Moreover, the AMC condition is slightly different between the two paradigms, with SanMiguel et al. (2013) including two separate scenarios for fulfilled and unfulfilled expectations. The sound always occurs at the same latency after the button press.

However, the findings from SanMiguel et al. (2013) do not separate the signal associated with prediction errors from other mechanisms of prediction processing, as it cannot be assumed that the EEG signal registered for unfulfilled predictions is exclusively related to prediction errors (SanMiguel et al., 2013). Several studies examine the neural activity of prediction errors by varying subtle features of the auditory stimuli that are more readily associated with variations in the electrophysiologic activity of the brain: for instance, it is possible to change the probability rate at which sounds appear, the sound properties, or the time delay of sound presentation. For example, Knolle et al. (2013) examined how the EEG signal related to predictions varies if the pitch of the sound presented after the button press is incongruent with the expectations. Altered sounds lead to a modulation of the self-suppression effect, which indicates the presence of a different prediction error signal from absent sounds, simultaneously detecting the error and the deviant stimulus (Knolle et al., 2013). Pinheiro et al. (2019) focused on the temporal dynamics bound to prediction errors, varying the delay to which sounds appear

after the action. The results show a reduction in the N1 suppression for sounds that arise after long delays; when the time delay between the action and the stimulus is significant, the brain might interpret the sound as externally generated (Pinheiro et al., 2019). Additionally, Alvaro D'Arriba and colleagues separated the electrophysiological correlates of sensory and intention-based predictions. The authors added some conditions to SanMiguel et al. (2013), presenting tones that can be congruent with sensory-based prediction, intention-based predictions, incongruent with both, or congruent with one while incongruent with the other. Their results replicate a general suppression effect for self-induced sounds; moreover, violated intention-based predictions are further differentially processed and correlate with an additional suppression effect of the P3b component reflecting the reappraisal of previous beliefs (Darriba et al., 2021). Overall, consistent research on predictive processing relies on the experimental paradigms just explained, but Horváth himself (2015) expressed some criticism about the methods (Horváth, 2015). First, studies using the Contingent paradigm extend the conclusions about predictive processing to findings based only on simple action, namely speech vocalizations (Arnal et al., 2011; Behroozmand & Larson, 2011; Wang et al., 2014) or button presses (Baess et al., 2011; Ford et al., 2014; SanMiguel et al., 2013). Alho et al. (2014) attempted to generalize the results by adapting the paradigm to other actions, but the current evidence is too weak to draw solid conclusions (Alho et al., 2014). Another crucial lack of consistency concerns the absence of changes in the rate of stimuli occurring after the actions (Horváth, 2015). Prediction errors are always present in the AOC and always zero in the AMC. Therefore, their probability is either 100% or 0%, without the possibility of examining ERP variations with changes in the rate of prediction errors. Few studies have also attempted to implement rate manipulations to vary prediction errors in magnitude and quality (Baess et al., 2011; Dogge et al., 2019; SanMiguel et al., 2013), providing evidence for sensory modulation beyond the simple detection of the stimulus. The current study will follow this line, presenting the sound with a 50% probability after the action to observe possible fluctuations in predictive processing. Moreover, the self-attenuation effect is multisensorial and multimodal; thus, controlling the variables that influence the formation of predictions using different conditions is highly relevant. Dogge et al. (2019) claimed a stronger suppression effect for self-cued stimuli, evidencing how the latter does not solely depend on action execution (Dogge et al., 2019). Harrison et al. (2021) introduced the role of two additional variables related to the timing of the stimuli, temporal predictability, and control. Temporal predictability is the possibility to predict the onset of the stimulus and must be internalised for self-generated precepts.

However, it is high for externally generated stimuli when cues about their occurrence are available. Temporal control is the feeling that a performed action can control the appearance of stimuli in a precise moment. When temporal control and predictability are equal among conditions, self-generated sounds elicit a reduced response compared to externally generated ones (Harrison et al., 2021). The results confirm the adaptive role of the suppression effect: the electrophysiological response to stimuli coming from the person decreases since there is no need to process their consequences, as stated by Ideomotor theories (Shin et al., 2010). The current study brings particular attention to this consideration, trying to separate motor and external cue-based predictions, growing from a more general predictive mechanism. The Auditory-Visual condition, in which visual cues predict the sounds, allows us to compare how predictive processing changes in this case. This way, the subject can make some inferences about the relationship between the two stimuli and create predictions based on external cues that, compared to motor predictions, might help to isolate its identity and the electrophysiological correlates related to it.

General considerations from the creation of the current paradigm. Several new hypotheses arise from revisiting the classic paradigms: predictions that generally affect the N1 amplitude without explicitly reflecting motor-based computations (Dogge et al., 2019; Harrison et al., 2021). Thus, we might expect a variation in N1 amplitude in both the Auditory-Motor and the Auditory-Visual conditions, carrying action-based and cue-based predictions. Moreover, if the N1 suppression builds on internal forward models (Miall & Wolpert, 1996), we should hypothesize a greater suppression effect for the AMC than for AVC (Bäss et al., 2008; Miall & Wolpert, 1996) due to the additive influence of self-generated actions in predicting upcoming stimuli. Therefore, we might expect a summative effect in the modulation of the N1 amplitude, with the predictive processing modality and prediction error magnitude being the factors of interest. On the other side, if the N1 component is related to cerebellar computations that are exclusively in charge of motor predictions (Knolle et al., 2013), the N1 might be affected by prediction errors during the Auditory Motor Condition only. As abovementioned, the current study also introduces a novelty in the investigation of predictive processing, analyzing the variations of the prediction errors from trial to trial. The classic version of the Contingent Paradigm only differentiates between conditions in which prediction errors appear with a 100% or a 0% probability (Horváth, 2015); in this case, the sound is either always present or absent. However, if predictions play a role in sensory attenuation, then the degree of prediction violation (i.e. the size of the prediction error) should also play a role. To see the influence of predictions on N1,

we varied prediction errors from trial to trial by presenting the sound only in 50% of them for every block. Then, we performed a single trial-based analysis, relating prediction errors to the amplitude of the N1 ERP component. This way, prediction errors should be present for the occurrence and the omission of the auditory stimulus since none of the cues work as a safe event. The statistical analyses will only focus on prediction errors calculated for the trials in which the sound is present, a new approach in predictive processing research.

Abstract

The ability to predict upcoming situations helps us to control the perceptual experience and keeps us from being overwhelmed by the complexity of the phenomenal world. Predictions can be extensive and diverse, involving cognitive, motor, and sensory elaborations; they are all studied within the theoretical framework of predictive processing, in which the brain is considered a sophisticated hypotheses-testing machine (Seth & Hohwy, 2021). Given the vastity of predictive computations, the current study focuses on the complex processing of auditory stimuli predicted by either self-generated or externally generated cues. Specifically, trial-by-trial variations of the N1 and the P2 ERP components during auditory predictive processing will be investigated, with the expectation of replicating the results of previous studies using the button-press paradigm (Knolle et al., 2013; SanMiguel et al., 2013; Schütz-Bosbach & Prinz, 2007). In addition, we will examine specific ERPs amplitude changes in the auditory-motor, the auditory-visual, and the auditory-only Conditions, namely the three experimental conditions in which the auditory stimulus is tied to different cues and appears only in 50% of the trials. Considering that the tone presentation rate is 50%, the ERPs amplitude might be modulated by the size of prediction errors in addition to the type of cue, with a possible summative effect of the two variables. The results of the mixed linear model analyses highlight a strong influence of the condition type and the size of prediction error on the P2 amplitude; however, there is no evidence for the suppression effect (Horváth, 2015) for both N1 and P2. Nevertheless, the study must be read with the inherent limitations of a pilot study in mind. Future research using the current paradigm should overcome methodological weaknesses that prevent a clear picture of the results, beginning by considering a larger sample and focusing on a more comprehensive examination of the behavioral data.

Methods

2.1 Participants

The sample comprises 20 healthy participants, selected among undergraduates from Heinrich Heine University (fourteen females, five males, one gender neutral, average age=22.55, SD=2.39; nineteen right-handed; one left-handed). The experiment was advertised with flyers and messages widespread among the university community and conducted inside the Biological Psychology department of the Heinrich Heine University in Düsseldorf, Germany. Normal or corrected to normal vision, normal hearing, no history of neurological disease and age between 18 and 35 years were the requirements for participation. For their participation, subjects received either exam credits or monetary compensation from 15 to 20 euros, depending on the total duration of the experiment.

2.2 Procedure

The current pilot study aimed to investigate the neural correlates of prediction error using a revised version of the paradigm designed by Seidel et al. (2021), which was a variant of the self-generation paradigm (SanMiguel et al., 2013). It consists of a within-subjects design with five experimental conditions and two training sessions, each displayed once and presented in separate blocks, each of which comprised 140 trials, for a total of 700. The presentation of blocks was randomized between participants. The choice of displaying each block only once allows for a smoother calculation of the prediction error indexes during data analysis, especially when analyzing only the trials in which the sound appears. This choice makes it easier to calculate the size of prediction error indexes and to pair those indexes with the corresponding EEG data. The procedure took place in an electrically and sound-shielded EEG room for better registration of the electrophysiological signal. Before starting with the experimental paradigm, participants signed the informed consent under the Declaration of Helsinki, ensuring the treatment and data protection. They also completed a computerized demographic questionnaire, the Edinburgh Handedness Inventory (Oldfield, 1971), and a paper form about the observance of the university rules related to the COVID-19 pandemic. Afterward, participants undergo two training sessions, always presented in the same order. Up to this point, no electrophysiological data were registered, but the acquisition of the EEG signal started at the beginning of the first experimental block. The total duration of the paradigm took from thirty-five to forty minutes, depending on how fast the subjects responded. The overall time of the session, including the

experimental setup and the EEG preparation, was approximately one hour and forty-five minutes.

2.3 Material

Questionnaires. Participants filled out a Demographic Questionnaire created in the Biological Psychology department of Heinrich Heine University. It allowed the collection of some general information: the date of birth, the biological and gender, the dominant hand, the history of neurological, medical, or psychological diseases, and information about alcohol and cigarette consumption. Such details are fundamental for verifying that the subject did not fulfill any exclusion criteria. Afterward, participants completed the computerized version of the 10-item Edinburgh Handedness Inventory (Oldfield, 1971), quantitatively assessing the hand dominance of the participant in activities of daily living (ADL). The items are in paragraph 5.3 of the Appendix; for each, the participant specified the hand or the side they prefer to perform the activity (self-rated method). When the preference is either similar or not present for a given activity, participants assign a “+” to each side; when it is for one side, they assign two “+” for that side; the item remains unmarked if the subject does not have experience with that specific task. The final score is the Laterality quotient, calculated via the formula reported in paragraph 5.3 Appendix. According to Oldfield (1971), the self-reported version is ubiquitous since participants tend to overestimate the frequency of use of their dominant hand; the version in which the participant got rated under direct observation should help overcome this problem. After the Inventory, participants answered three questions about the handedness of their parents.

Setup and tools. The study was designed and displayed via Presentation® (Version 20.3, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) on a BenQ (EW2740L, 27-inch LED, full HD) computer screen with a screen resolution of 1920 × 1080 and a 60-Hz refresh rate; Presentation® was also used to send the ERP markers. The first training was created and displayed separately via PsychoPy 2020.2.10, a software for running a wide range of experiments in behavioral science (www.psychopy.org/). For the conditions requiring button presses, participants used a Cedrus RB-740 response pad (www.cedrus.com); this box-shaped tool presents seven response buttons along a line, set to interact with the displayed paradigm. As illustrated in Figure 2.1, the response box was positioned between the monitor and the subject. Participants only need to use the central button for the task and the leftmost button to go from one block to the following. Moreover, participants wore over-ear headphones (Sennheiser HD 201)

for the auditory stimulus delivery. While the participant completed the paradigm, the experimenter was in the control room, monitoring the EEG recording and the correct performance of the procedure through two BenQ computer screens (EW2740L, 27-inch LED, full HD).

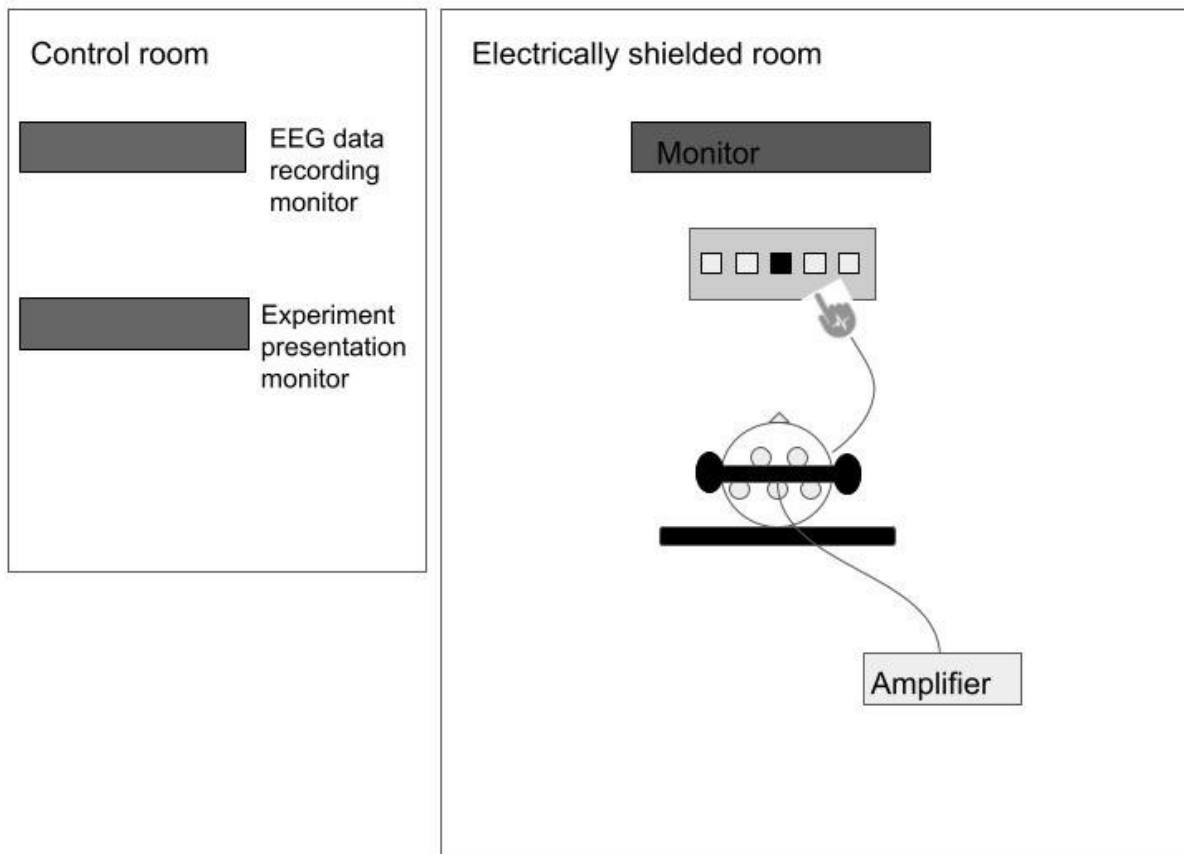


Figure 2.1 Experimental Setup. The image is an adaptation of Figure 1 of the paper by Ghio et al. (2018)

Stimuli. Figure 2.2 shows the visual stimuli presented during the paradigm. These stimuli were presented on a black background, which was also maintained for the instructions. The left icon, consisting of a button press symbol, was only present during Training 1. The second is a solid blue circle with a white fixation cross in the middle, appearing during the second training, the Auditory-Visual Condition (AVC) and the Visual Only Condition (VOC). The blue circle, created using Presentation® software, is an ellipse graphic with a font size of 150, while the fixation cross is the text element “+” with a font size of 40. The fixation cross remains on the screen for the entire duration of the paradigm, helping the participant to keep the focus of attention on the monitor. The auditory stimulus appears in each block beside the Motor Only

Condition (MOC) and the Visual Only Condition (VOC) with a 50% probability of occurrence. It was a 680Hz sinus tone of 50 ms and a button press-onset latency time of 300 ms, delivered binaurally and with equal intensity for all the participants.

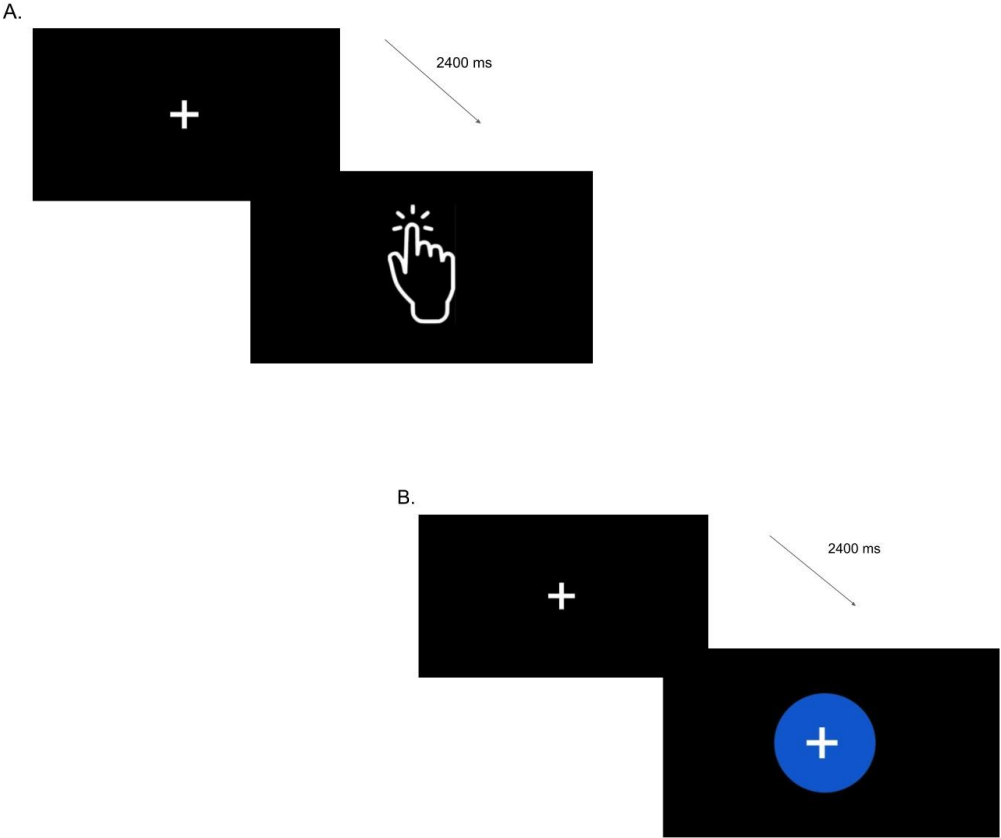


Figure 2.2 Visual stimuli displayed during the study. Figure A shows the stimulus used in Training 1; Figure B shows the stimulus delivered in Training 2, the AVC, and the VOC. Both stimuli appeared at the rhythm of 2400 ms (with 200 ms of random variance).

2.4 Conditions

Training 1. During the first training session, participants had to press the button on the response pad every time the button press symbol appeared on the screen. The sound followed the action only in 50% of the trials. The purpose is to make the participant learn the action rhythm, consisting of a button press every 2400 milliseconds (\pm 200 milliseconds of random variance as the window tolerated for data analysis). In this case, it is not the tone that provides the rhythm for the button press but the display of a button press icon: this allows to avoid the association between the sound and the rhythm and thus predictions that link the two events. During the instructions, the participants are warned to memorize the rhythm without counting or using any cognitive strategies to avoid introducing unrelated computations in the experimental procedure.

Training 2. The second training, identical to the Auditory-Visual condition (AVC), enables the participant to create an association between the visual cue and the upcoming sound, forming cue-based predictions that do not depend on actions. The participant only needed to watch the monitor without performing any action. The fixation cross was always displayed, while a blue circle appeared on the screen at fixed intervals of 2400 ms between one visual stimulus and the following; as in the first training, the sound accompanied the blue circle in 50% of trials only.

Auditory-Motor Condition (AMC). During the Auditory Motor Condition (AMC), the subject pressed the response pad button at the same rhythm learned during the first training. The screen only displays the fixation cross for the whole duration of the block, and no visual stimuli are present. In contrast to the AMC condition of the classic self-generation paradigm, in which each button press elicits a sound (Horváth, 2015), only 50% of actions were followed by the sound; in this case, the sound appeared with a latency of 300 ms after the button press. The AMC allows the formation of predictions regarding self-generated stimuli since the performance of the action is linked with the appearance of the sound. Since the sound follows the button press in half of the trials only, this condition carries prediction error computations for the trials in which the sound is absent. Most importantly, prediction errors are present also in the trials in which the sound is present since the latter is not expected with a 100% of probability; this is the most important point since only those trials will be later analyzed.

Auditory-Only Condition (AOC). In the Auditory Only Condition (AOC), the participant is instructed to listen to the sounds delivered through the headphones at the same regular interval as before, namely 2400 plus 200 ms of random variance. There is no need to press the button or perform any other action during the block. The AOC allows registering the EEG signal of

the activity related to the sensory elaboration of external sounds, with no possibility of control.

Motor-Only Condition (MOC). The Motor Only condition (MOC) allows for the registration of the electrophysiological responses related to motor execution only. Here, the participant needed to press the button at the learned rhythm of 2400 milliseconds per button press without any sound occurring; the fixation cross is the only stimulus displayed on the monitor.

Auditory-Visual Condition (AVC). During the Auditory-Visual condition (AVC), the participant did not have to press any button but just had to look at the screen. In this case, the blue circle appeared on the monitor at fixed intervals of 2400 ms between one stimulus and the following, analogously to the time rhythm learned for the button press. The sound accompanied the blue circle in half of the trials and appeared with a latency of 300 ms after the visual cue. The cognitive elaborations relate to cue-induced predictions, namely predictions generated for stimuli coming after simple cues, such as the blue circle, but no actions.

Visual-Only Condition (VOC). Analogously to the AVC, in the Visual Only Condition (VOC) no button press is required, and the participant passively observed the monitor. The blue circle appeared in the middle of the screen at a rate of 2400 milliseconds between stimuli; conversely to the AVC, no sound accompanied the blue circle during the block. This condition allows for extracting the activity related to visual processing only from that registered during the AVC, during which predictions upon the visual stimulus and the sound arose. As for the MOC, when subtracting the VOC from the AVC, it is possible to obtain a corrected AVC signal, which relates to predictive processing computation concerning the association between the visual cue and the sound.

Overall, the experiment presents five conditions and two training sessions. The training sessions aim to create associations between the auditory stimulus and different predictive events before the beginning of the experimental blocks. The association of the visual cue with the sound is fundamental to controlling the formation of cue-induced predictions between the two stimuli. Concerning the experimental conditions, the AMC and the AVC unfold the elaboration of different prediction error signals. The AMC relates to PE based on prediction from own movements, while the AVC unfolds PE arising from environmental-related predictions. The VOC and the AOC allow registering the activity linked to pure visual and auditory

computations, serving as control conditions. The MOC disentangles the activity related to motor processing; it makes it possible to obtain a motor-corrected AMC signal, carrying only the activity present during predictive processing. Moreover, participants used their right hand for the button press in every condition, besides their hand dominance, to reduce within-subject variability. Participants could take a break between conditions and voluntarily control when starting with the following block by pressing the left button on the keyboard; this allows them to reduce their tiredness and optimize the employment of their attentional resources during the task.

2.5 EEG data acquisition and analysis

Data acquisition. The EEG signal was continuously recorded after the two training sessions, starting from the first experimental condition. It was registered at a sampling rate of 1000 Hz via the Brain Vision Recorder software (1.20.0506, Brain Products, GmbH, Germany), and acquired by applying twenty-eight Ag/AgCl passive ring electrodes on the scalp, inserted within an elastic cap (www.easycap.de). The BrainAmp Standard amplifier (www.brainvision.com) allowed for signal amplification; the impedance was kept below five kOhm. The montage was monopolar, while the placement of electrodes followed the rules of the International 10-20 system (Luck, 2014): scalp positions 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, and PO8 were the ones used. Consistently with Knolle et al. (2013), the average activity from the mastoids was used as reference and registered by placing two additional electrodes at TP9 and TP10, while the ground electrode was at FCz. To be able to control the electrical activity coming from vertical eye movements, one electrode was positioned below the right eye. The latter aligns with the electrode in Fp2; the two taken together register blinks and other vertical eye movements. For the same reason, the electrooculogram was tracked from electrodes F9 and F10 to register horizontal eye movements. More information about the Electroencephalography montage and basics are in the Appendix, paragraph 5.1.

Data analysis. Brain Vision Analyzer 2.2 (Brain Products), MATLAB (R2018a, The MathWorks, Inc., Natick, MA), and R studio (version 1.2.5019) are the software packages used for pre-processing and analysis of EEG data. The participants for whom less than 70% of trials for at least one of the five experimental conditions could remain in the analyses were excluded. This has been the case for three participants: therefore, 17 participants (fourteen females, three males, and one other, average age=22.68 SD=2.35; sixteen right-handed, one left-handed) constitute the final sample.

Pre-processing. Pre-processing was implemented on Brain Vision Analyzer 2.2 and started with a Butterworth zero-phase filter (low cutoff: 0.3 Hz, 12 dB/oct; high cutoff: 30 Hz, 12 dB/oct) and a Notch filter (50 Hz) to remove current-related artifacts. The Independent Component Analysis (ICA steps = 512, Infomax restricted biased) extracted the statistically independent components of the signal; an inverse ICA in semi-automatic mode followed the latter. The resulting statistically independent compounds underwent visual screening to detect

and remove the ones reflecting eye artifacts, which usually mirror the typical electrophysiological pattern for eye blinks. These components were selected and the electrode sources originating them were deactivated. Successively, the resulting signal was segmented by time-locking the ERPs of interest to the sound markers, namely the labels used to indicate the presence (or absence) of a sound event. After this procedure, the final segmented epochs have a total length of 800 ms, starting at 200 ms before the sound and ending 600 ms after. When time-locking the ERPs in the AMC and the AVC to the sound markers, only the ones indicating the tone appearance were considered. Indeed, the following analyses on prediction errors focused on the fifty percent of trials in which the sound appears. It was also possible to time-lock the “no-sound” events to the ERPs of interest: when no sound followed the button press, the tone was indeed still presented but muted. Thus, even though no sound was present, the time-locking to an event was still possible; also, the first trials of each block and the ones in which the button presses were too fast or too slow were coded with different markers and not included in the segmentation. All the segments were baseline corrected, considering the mean amplitude of the signal within 200 ms before the muted tone as the baseline signal. The ones containing artifacts were detected and deleted from the analyses using the automatic algorithm of Brain Vision Analyzer 2.2 (maximally allowed voltage step = 50 $\mu\text{V}/\text{ms}$, a maximally allowed difference within 100-ms intervals = 100 μV , maximal/minimal allowed amplitude = ± 100 μV , lowest activity of 0.5 μV within 100 ms intervals). According to the standard procedure of the Contingent Paradigm (Horváth, 2015), we calculated the individual motor-corrected AMC waveform by computing the differential wave between the single subject average MOC signals and the AMC signals and allowed to delete pure motor computation from the latter; analogously, the individual average VOC signal was subtracted from the AVC signal to obtain the visual-corrected AVC waves. The grand averaged signal ERP components before and after such correction are illustrated in Figures 2.3 (a) and 2.3 (b). For simplicity, the acronyms AMC and AVC refer to the motor and visually correct signals from this point.

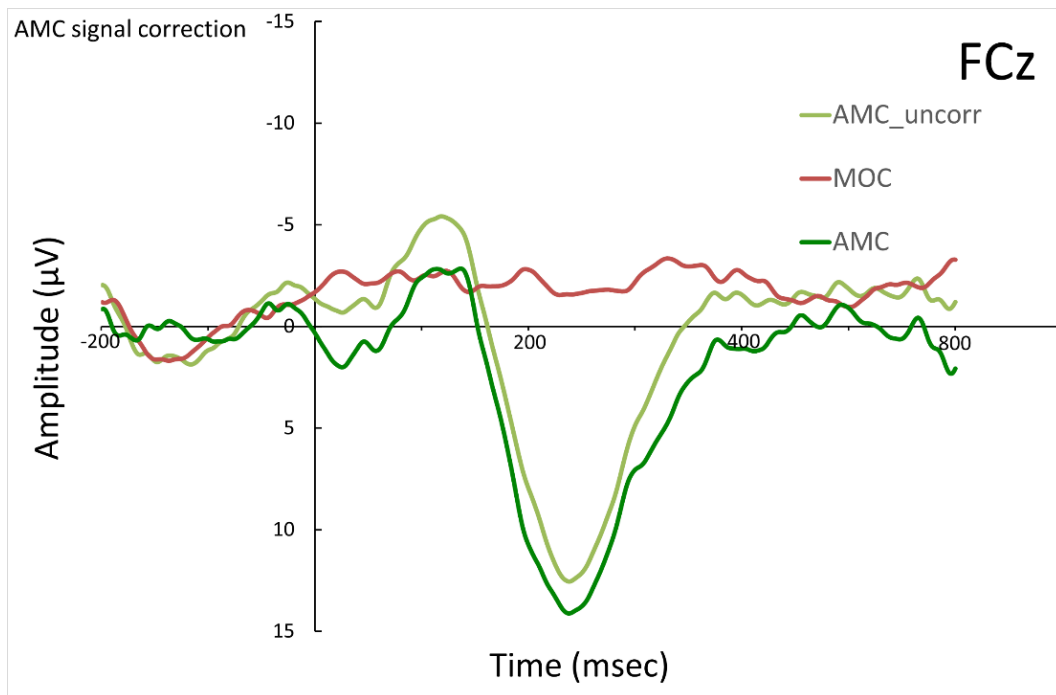


Figure 2.3 (a) The visual correction of the AMC is displayed at electrode FCz. The corrected AMC waveform is obtained by subtracting the ERPs for the Motor-only condition (red) from the ERPs associated with the uncorrected AMC signal (light green)

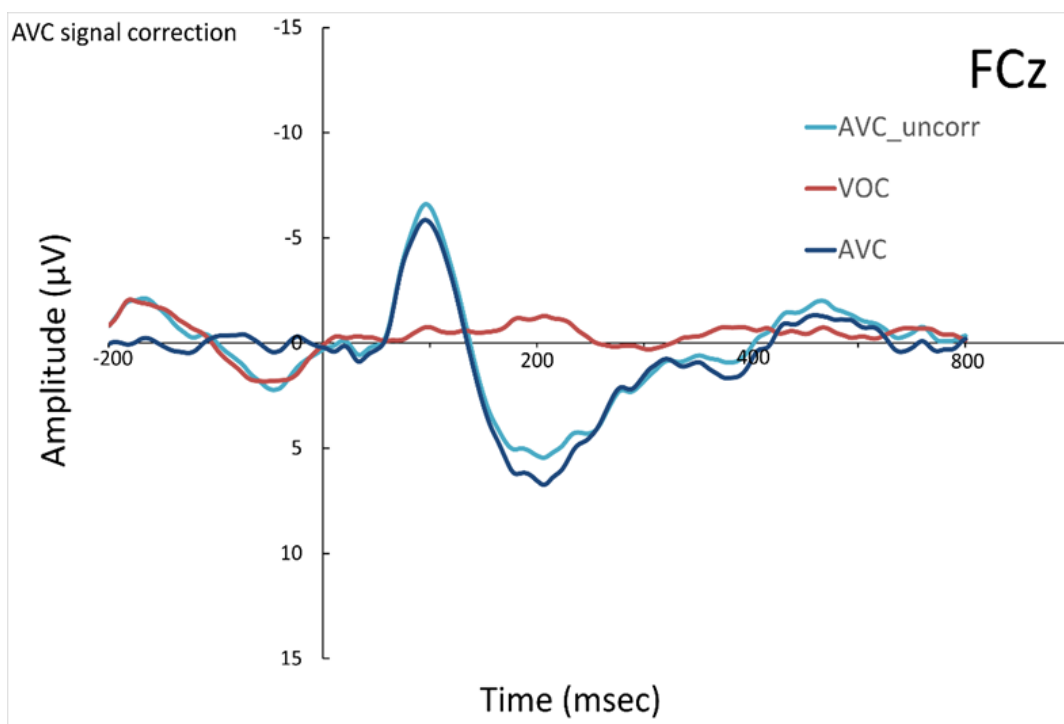


Figure 2.3 (b) The visual correction of the AVC is displayed at electrode FCz. The corrected AVC waveform (dark blue) is obtained by subtracting the ERPs for the Visual-Only condition (red) from that of the uncorrected AVC signal (light blue).

Computation of grand averages and topographies. Second, we used Brain Vision Analyzer to compute the grand-averaged distribution of N1 and P2 for AMC, AVC, and VOC from the primary history files, selecting the individual averaged signals for each condition. The ERPs were detectable and lent themselves to a preliminary visual analysis. Indeed, the visual inspection of the topographical distributions (Figure 3.2) and the grand-averaged waveforms (Figure 3.1) helped set an adequate time window for the extraction of N1 and P2 peaks, fundamental for the statistical analysis and usually going from 100 ms to 200 ms and from 200 ms to 300 ms respectively. The topographies of N1 and P2 were computed on the signal component at the electrode FCz, choosing the display of four maps and manually scaling the time window of interest for each topography. Each of the four maps embraces different time windows to show the signal course of the component in a precise way. The grand-averaged ERPs were computed at FCz, Cz, and Fz, the three electrodes included in the statistical analyses. Following the guidelines from previous studies about auditory predictive processing (Ghio et al., 2018; Seidel et al., 2021), we considered only the signal from these electrodes because of the greater involvement of frontoparietal areas during the button press paradigm (Horváth, 2015).

Calculation of PE. Matlab was used to calculate the theoretical PE indexes for each trial based on the rate at which the sound is present and the consecutive presentation of the ones in which the sound appears. The formula for the calculation derives from the theoretical account of Friston et al. (see paragraph 5.2 of the Appendix), stating that *prediction error = external stimulus – expectations*, with an initial expectation value set at 0.5. At the beginning of the task, the participant does not have any expectations about the tone presentation; at this point, the expectation index can assume a value of 0.5. The upcoming expectation values arise by the formula: *expectation = expectation start (prediction error * α)*, considering the index α as the learning rate index with a fixed value of 0.2. However, the alpha value has been arbitrarily set and can be varied to see how the data change using different values. Once the PE and expectation values were calculated we bound the data to each block and associated them with the electrophysiological data. Second, we sorted the grand-averaged ERPs data into three groups based on the PE magnitude. The output is a table with three columns named low, medium, and high PE. The medium PE column encloses the data that are $\pm \frac{1}{2}$ SD around the mean, while the low and high columns consider the data below and above the range of the medium column. From the data in the table, we created Figure 3.3 (see Results section),

showing the grand-averaged ERPs for different PE magnitudes, and underpinning how the latter might influence their amplitude. This computation has been done on the signal component at FCz only.

Statistical analyses. Matlab allowed the extraction of the single-trial N1 and P2 average peak values for each subject by considering their average amplitudes within the time window of 40 milliseconds from the maximum one and setting a standard deviation value of 2.5 for the detection of the outliers. Then, the peaks were associated with the individual PE and expectation values to conduct a trial-by-trial statistical analysis. The statistical analyses were performed in the R environment, separately considering N1 and P2. For all the inferential statistics, an alpha threshold of 0.05 was stated for the presence of significant results. Several separate Mixed Linear Models (MLM) (Meteyard & Davies, 2020) were performed using the *lme4* and the *lmerTest* packages (under R version 4.2.1). This approach allows defining models in which one or more predictors interact to predict the outcome variables, namely the amplitude changes of N1 and P2. Specifically, we created two Mixed Linear Model analyses, separately run for each ERP. The first MLM included the categorical variable Sound type, simple coded and considered as the fixed effect predictor (-1/3; -1/3=AOC, 2/3; -1/3=AVC, -1/3; 2/3=AMC), and the factor subjects as the random-effect predictor; as stated above, the amplitude of the frontocentral ERP components is the dependent variable of the model. Then, we investigated the effect of PE in predicting the ERPs amplitude, including this variable in a second MLM performed for both N1 and P2. We defined the model considering the categorical variable Sound type without the level AOC, which should not be affected by the PE magnitude; therefore, the Sound type variable was anon simple coded as a two-level variable (0.5=AVC, -0.5=AVC). The continuous variable PE was mean-centered; the random-effect variable is the same as the one described above; further simple effect analyses studied the interactions between Sound type and PE. Two supplementary MLMs considered low and high PE values only by shifting the center of PE of one standard deviation up or down. In addition, we tested each MLM with an ANOVA test performed with Satterthwaite's method to detect their overall predictability. The package *ggeffect* (under R version 4.2.1) calculated the predicted values of the dependent variables at the margin of specific model terms, with the confidence interval set at 0.95; pairwise comparisons between variables were computed using the function *emmeans* (under R version 4.2.1).

Results

3.1 ERPs

N1 Component. The first Mixed Linear Model testing the three-level Sound type factor as the only predictor of N1 amplitude revealed no significant main effect of the latter in predicting any dependent variable change, $F(2,16) = 0.91, p > 0.05$. The second Mixed Linear Model, excluding the AOC from the Sound type levels and including the factor PE within the fixed predictors, highlighted no significant main effect of the independent variables or their interaction (all $p > 0.05$).

P2 Component. As described in the Methods, the Mixed Linear Models designed for the P2 analysis were analogous to the ones created for the N1 analysis. P2 peaks at a different latency in the AMC as compared to the AVC and the AOC (see Figure 3.1); thus, the average amplitudes of P2 were extracted from a different time window in this condition, automatically selected via Matlab. The first Mixed Linear Model revealed a significant main effect of the three-level categorical factor Sound type $F(2,16) = 12.07, p < 0.05$. The average P2 amplitude significantly differed in the AMC compared to the AOC $t(16) = 4.78, p < 0.05$ and the AVC $t(16) = 3.29, p < 0.05$, with the P2 amplitude for the AMC ($M = 14.03, 95\% \text{ CI} [11.06, 16.99]$) being significantly increased as regards to the AOC ($M = 6.79, 95\% \text{ CI} [5.47, 8.10]$) and the AVC ($M = 7.83, 95\% \text{ CI} [4.15, 11.51]$). Table 2 shows the parameter-specific values for this model.

The second Mixed Linear model, built on the analogous independent variables which are present in the second model for N1 analysis, revealed a significant effect of PE $F(1,16) = 19.68, p < 0.05$, Sound type $F(1, 16) = 11, p = 0.004$, and PE x Sound type $F(1, 14.87) = 8.29, p = 0.01$, with PE being the strongest predictor of P2 amplitude changes. Pairwise comparisons revealed a significant P2 amplitude increase for PE x AMC ($M = 14.07, 95\% \text{ CI} [10.8, 17.3]$) compared to PE x AVC ($M = 7.79, 95\% \text{ CI} [3.8, 11.8], t(16) = 3.32, p = 0.004$). The interactions were further explored by performing two additional Mixed Linear models, which allowed us to separately test the effects of PE x AMC and PE x AVC. By looking at the specific contrasts, no main effect of PE was revealed by the model testing PE x AVC $t(16) = 1.7, p = 0.1$, even though its main effect was found to be strong for the model testing PE x AMC $t(16) = 5.08, p < 0.001$. Table 3 reports the parameter-specific values for the second MLM. Last, we

designed two MLMs to separately test the effect of high and low PE values on P2 (see Methods section for more details on the calculation of high and low PE values). Whereas PE and PE x Sound type had the same statistical power they assumed in the general PE x Sound type model, Sound type had a stronger effect for the model testing high PE values $F(1, 16) = 12.56, p < 0.05$. In addition, its main effect was reduced, but still significant, for the model considering low PE values $F(1,16) = 7.89, p = 0.01$.

Descriptive statistics. R studio was used to calculate the descriptive statistics reported in Table 1, considering the raw single-subject averaged ERP amplitudes. More specifically, the table shows the mean and the standard deviation of N1 and P2 amplitudes for each experimental condition and each electrode of interest.

Table 1. Descriptive statistics related to the dependent variables.

Electrode		AMC		AVC		AOC	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
N1 analysis	FCz	-3.77	10.66	-5.62	10.66	-4.66	9.83
	Fz	-3.76	10.99	-5.32	11.55	-4.77	10.02
	Cz	-3.67	10.26	-4.89	9.84	-4.51	9.51
P2 analysis	FCz	15.18	14.40	7.73	13.06	6.72	12.50
	Fz	15.89	15.19	6.81	14.25	6.04	13.00
	Cz	12.83	13.12	7.28	11.38	5.90	11.80

Table 2. Summary of the estimated effects from the first MLM testing P2. Sound type is the only independent variable, and AOC is set as the reference level. By using simple coding, the intercept reports the dependent variable effects computed on the grand mean of data.

Sound type. f2 shows the contrasts between AOC and AMC; Sound type. f3 shows the contrast between AOC and AVC; Sound type. f4 reports the contrasts between AVC and AMC, calculated by setting AVC as the reference level.

	Estimate	Standard Error	df	t-value	p-value
MLM Sound type (AOC, AVC, AMC)					
(Intercept)	9.55	1.05	16.07	9.06	<0.001
Sound type. f2	7.24	1.50	16.17	4.80	<0.001
Sound type. f3	1.04	1.73	16.00	0.60	0.55
Sound type. f4	6.20	1.88	16.14	3.30	0.004

Table 3. Summary of the estimated Mixed Linear Model effects on P2 amplitude changes considering Sound type (AMC, AVC) x PE. The reference level is AVC.

	Estimate	Standard Error	df	t-value	p-value
MLM Sound type (AVC, AMC) x PE					
(Intercept)	10.39	1.43	16.07	7.67	<0.001
PE	9.03	2.04	16.93	4.43	<0.001
Sound type	6.28	1.89	16.13	3.32	0.004
PE: Sound type	9.50	3.30	14.87	2.88	0.01

3.2 Grand averages and topographies

Figures 3.1 and 3.3 show the grand averaged ERP waveforms for the AMC (green line), the AVC (blue line), and the AOC (grey dashed line). Figure 3.2 shows the topographies of the N1 and P2 at the electrode FCz when the latter are time-locked to the sound markers.

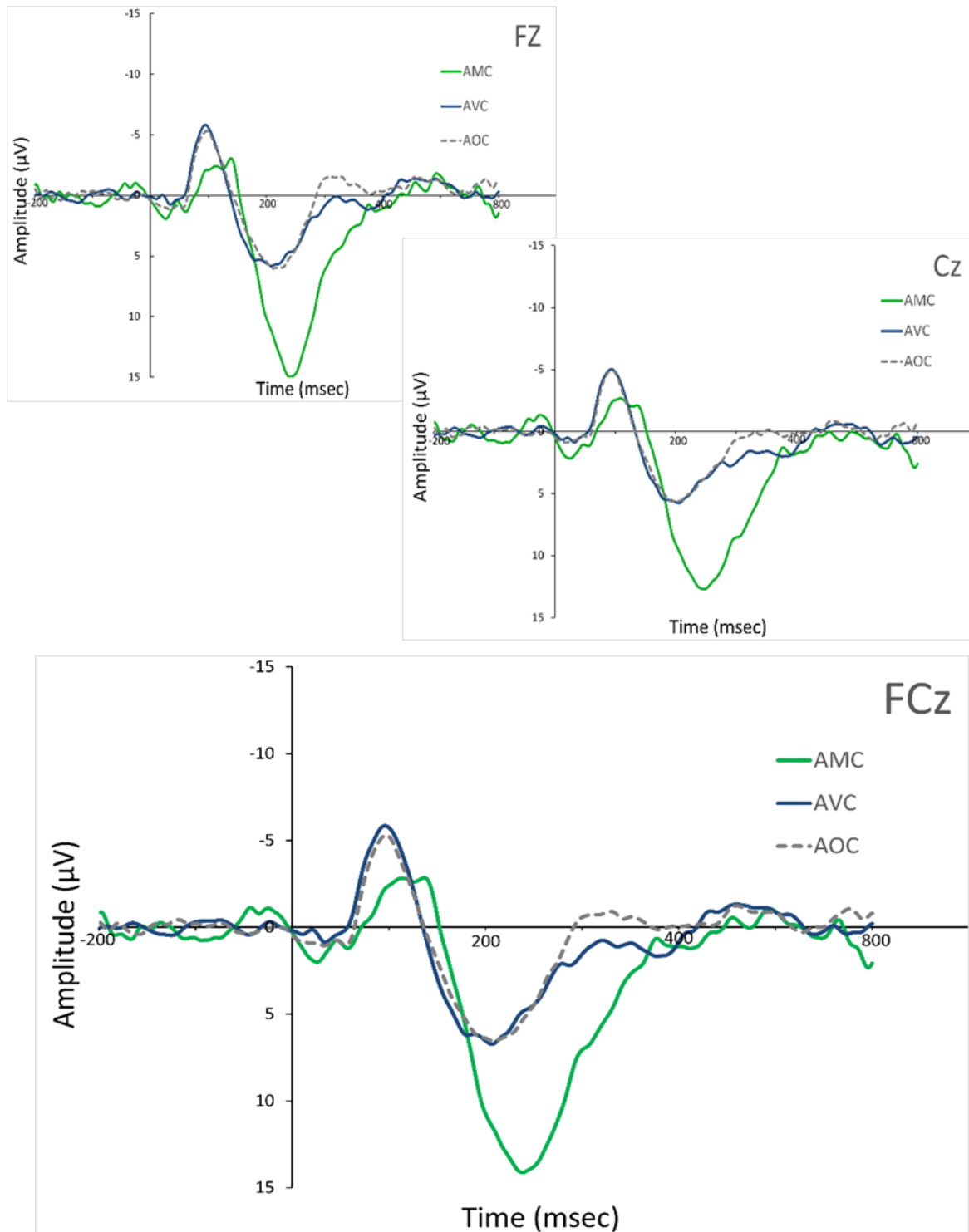


Figure 3.1 Grand Averages ERPs for Fz, Cz (small layout), and FCz (big layout).

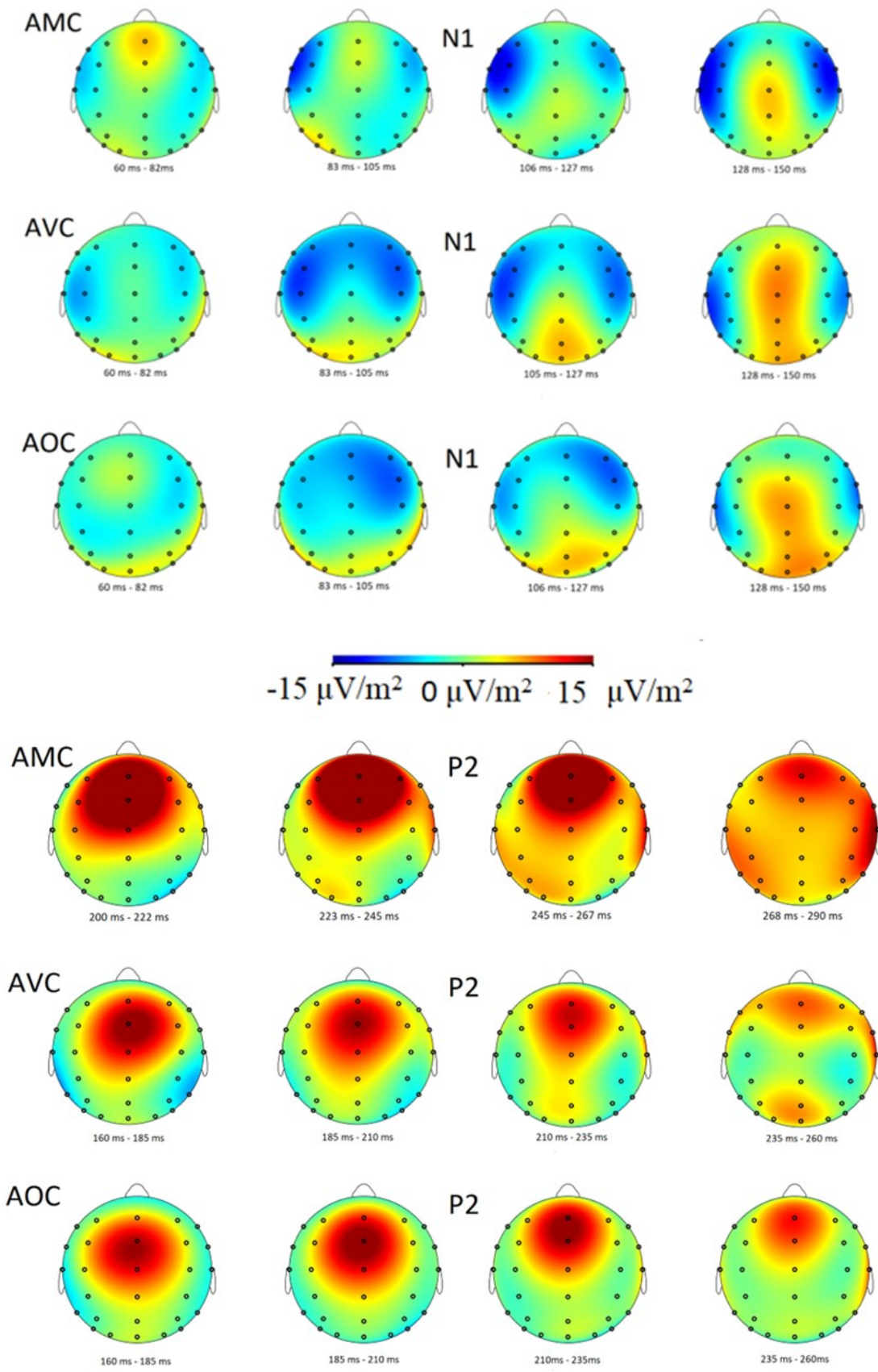


Figure 3.2 Topographies for N1 and P2 at the electrode FCz.

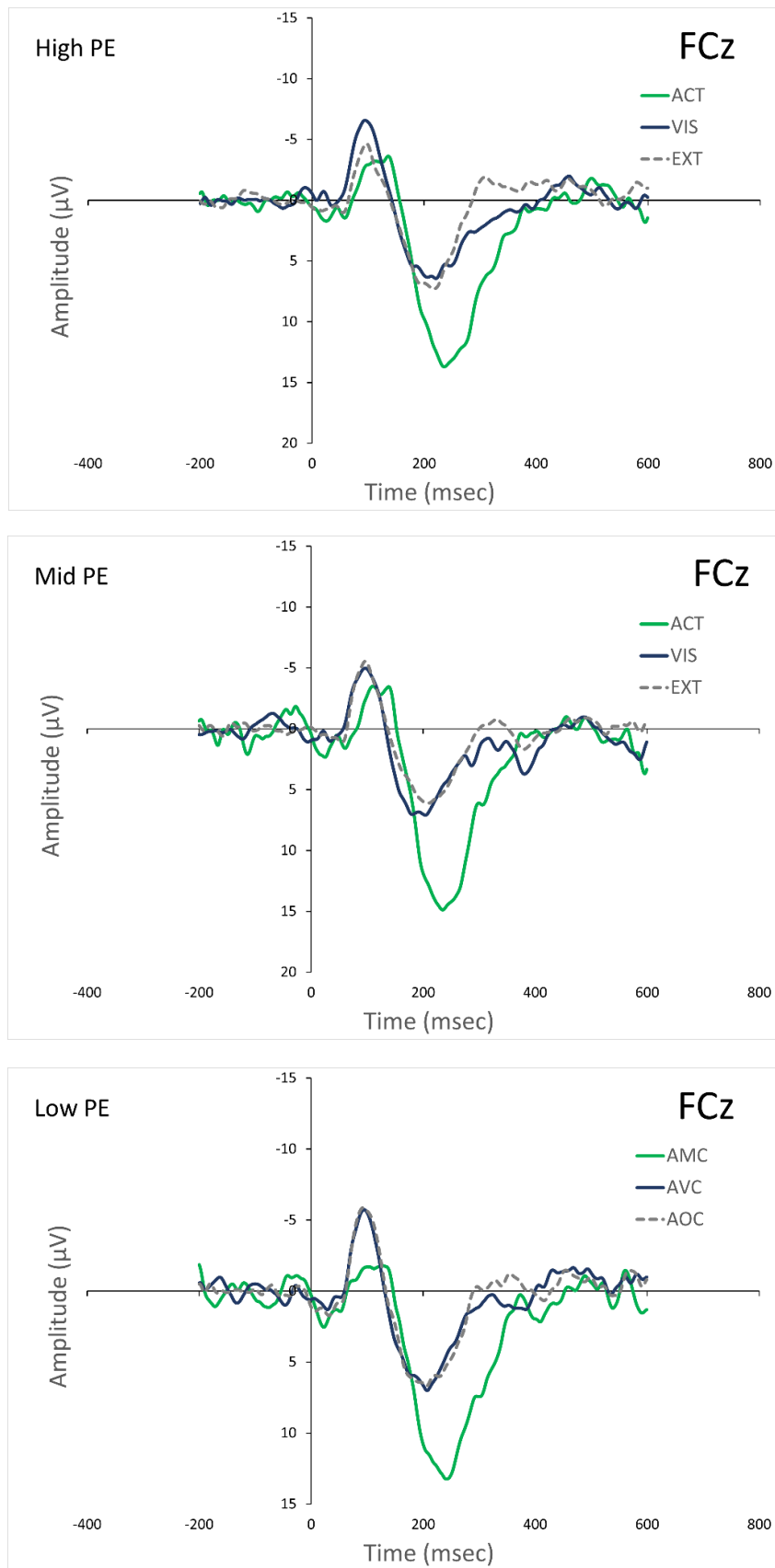


Figure 3.3 Grand - averaged ERPs according to PE magnitude levels (low, medium, high).

Discussion

4.1 Interpretation of the results

This pilot study aimed to investigate the within-subject trial-by-trial variations of prediction error electrophysiological correlates bounded to self-induced and externally generated cues when the target auditory stimulus is present in 50% of the trials. It does not replicate the findings about the suppression effect (Horváth, 2015; SanMiguel et al., 2013; Sowman et al., 2012), claiming an overall N1 and P2 reduction during the predictive processing of self-initiated auditory stimuli.

Regarding N1, no significant results arise from the statistical analyses related to it; therefore, the following considerations attempt to give meaning to the data, keeping in mind that is not possible to infer any clear conclusion. On these premises, the N1 amplitude is reduced in the AMC (see Table 1, Results) compared to the AVC and the AOC, increasing the hypothesis of a suppression effect.; further studies with a larger sample of participants can help clarify if the current data pattern is only due to chance. Second, the average N1 amplitude is almost equivalent in the AVC and the AOC (see Table 1 and Figure 3.1, Results). Thus, the hypothesis in aid of a general predictive processing mechanism (Dogge et al., 2019; Harrison et al., 2021) should be discarded, reconsidering those models that claim the presence of different processing paths for self and externally cued stimuli. For this purpose, theories relating the suppression effect to cerebellar motor predictions (Knolle et al., 2013) might better suit the interpretation of the results: they link the N1 suppression to the occurrence of specific motor-based inferences computed by the cerebellum, excluding a possible correlation between N1 changes and predictive processing related to other sensory modalities. Observing the course of N1 for the AVC and the AOC in Figure 3.1, we can see how the grand averaged waveforms are almost overlapping in these experimental conditions, proving the ineffectiveness of visual cues in predicting N1 variability. Therefore, a more definitive study using the current paradigm should support this interpretation rather than giving credit to those theories stating the N1 involvement in general predictive computations relies on the thalamocortical cerebellar pathways (Bansal et al., 2018; Bastos et al., 2012; Dogge et al., 2019; Sowman et al., 2012).

Sound type, PE, and Sound type x PE are all significant predictors of P2 amplitude modifications. The Mixed linear model setting Sound type as the only fixed effect on P2 reports a significant increase of the latter in the AMC compared to the other experimental conditions. In addition, no significant P2 amplitude differences between AVC and AOC have been found

(see Results section, Table 2), as visible from the grand averaged waveforms reported in Figure 3.1. Several studies investigating auditory predictive processing with the button press paradigm have reported a P2 suppression when the perceived senses of control and agency over the ongoing task are high (Moore et al., 2009; Thompson et al., 2007). In this case, participants can hear the sound only in 50% of the trials, which makes it difficult for them to have a perception of control over its presentation by pressing the button during the AMC; indeed, when the appearance of stimuli is unpredictable after the action, the amplitude of P2 is found to be increased (Pinheiro et al., 2019). According to previous literature, a P2 increment also manifest during the occurrence of those adaptive mechanisms leading to an update of the internal model cognitive model for predictions, which is essential when the expectations about the tone are unfulfilled (Behroozmand & Larson, 2011; Sowman et al., 2012). Such computations are more evident by studying the trial-by-trial variations of P2, underlying the ongoing predictive processes during each condition. As observable in Figures 3.1, 3.2, and 3.3 (see Results), P2 also peaks with an average delay of 60 ms in the AMC compared to the other levels of Sound type. A delayed presentation of this ERP component might be an early activation of the P3a, which typically peaks after 300 ms from the stimulus (Luck, 2005) and is functionally linked to P2 (Seidel et al., 2021). Hence, it can be more appropriate to talk about a P2/3a complex instead of considering the P2 component only. The up-to-date literature describes the frontocentral P2/3a as a neural signature of working memory and attentional mechanisms, associating its increase with the appearance of previously learned and expected stimuli (Lenartowicz et al., 2010). In the current paradigm, the detection of the tone might be influenced by Training 1, during which participants learn the button press rhythm (and its association with the possible appearance of a Sound) by repeating the action every time an intermittent button press icon was presented on the screen (Figure 2.2 A). Since the button press rhythm needs to be retrieved and repeated during the AMC, participants activate the working memory template for this information created during Training 1, comparing the current experience to the stored one. Because of the possible effect of the training on the results, a future study excluding it from the paradigm can be useful to understand their actual influence on the two ERPs under investigation. Additionally, the attentional resources of the subject are towards the sound, whose detection increases the amplitude of the P2/3a complex (Lenartowicz et al., 2010), especially when the participant has the feeling of not having control over its presentation. To our knowledge, this type of process is more functionally bound to the P3a than the P2, which is frontally maximal and reflects the allocation of automatic attentional processes (Lenartowicz

et al., 2010; Luck, 2005). Future studies might help to disentangle the functional differences between the latter and P2 by time-locking the P3a to the tone.

As aforementioned, P2 amplitude does not statistically differ between the AVC and the AOC. Consequently, P2 is not influenced by visual cues when the tone only appears in 50% of the trials. The blue circles, our visual cues, might be elaborated as external stimuli without being bound to any prediction. Nonetheless, Sowman and colleagues (2012) report a partial P2 suppression during the predictive processing of visually cued stimuli. Changing the sound rate might help to understand the results, to test whether an increase in the sound presentation rate leads the participant to interpret the visual cue as a predictor. In addition, P2 amplitude varies when considering different sensory modalities and stimuli (Mifsud and Whitford, 2017); hence, the current dependent variables might not fully capture the neural dynamics of visually cued predictions. Studying the signal components from different electrodes or performing a whole brain analysis might help to reach this goal. Furthermore, the magnitude of prediction error indexes positively correlates with a P2 amplitude increase, but it does not predict P2 changes in the model testing AVC x PE; therefore, prediction errors only affect P2 when predictive processing relies on self-generated actions. The findings support the hypothesis that visual cues are processed as external stimuli without any prediction error signal during the AVC. Additionally, PE magnitude positively correlates with the Sound type effect on P2 (see Results, last paragraph); hence, higher PE values increase the contrast between the Auditory Motor and the Auditory-Visual condition, leading to a higher increase in the P2 during the AMC.

Overall, P2 changes have been more related to general high-level computations, hardly explained by internal forward models. The up-to-date knowledge about auditory predictive processing states that the association between the self-generated cue and the occurring stimulus also bounds to some general interpretations of the context (Horváth, 2015). Therefore, N1 and P2 correlate with central computations rather than peripheral, mostly engaging frontocentral and frontoparietal areas. For this reason, the electrophysiological components examined come from the electrodes FCz, Fz, and Cz. Indeed, when the auditory stimuli appear after short intervals from the button press (within 4 seconds from the action), N1 has a major frontal or frontocentral distribution (Horváth, 2015). Regarding P2, this ERP originates from different sources because of its link to different computational processes (Luck, 2005); posthoc analysis considering signal components besides the ones from the frontocentral electrodes might reveal its presence in other areas, possibly related to computations other than the ones investigated. Therefore, the absence of consistent P2 variations in the AVC is more likely to rely on the

impossibility of starting predictive elaborations due to the inconsistency of the sound presentation and not its specific involvement in motor computations.

Last, predictive processing computations have been correlated to different personality traits or further experimental manipulations. For this purpose, some studies administer questionnaires or include training sessions in the experimental paradigm. To make some examples, Schafer and Marcus (1973) report a significant association between the suppression effect and the IQ Intelligence scores of the participants (Schafer & Marcus, 1973). Seidel and colleagues (2021) have used a very similar paradigm to the one created by SanMiguel (2013) while manipulating the Illusion of Control (IoC) in a training condition (Seidel et al., 2021) since the perception of being highly responsible for action consequences is highly influential in shaping predictions (Ford et al., 2013; Kühn et al., 2011). Empathy scores might also impact predictive processing: this correlation emerges from ERPs variations for externally generated sounds or sounds induced by another person. When the experimental subject is just a passive observer, it is possible to investigate how the electrophysiological response differs from that of the performer and investigate if the observer can get into the picture of the performer itself (Angelini et al., 2018). However, discordant findings about the influence of these manipulations have emerged. Despite all the possible variations, the present study delves into the classic mechanisms of predictive processing, focusing on the pure computation of prediction errors from a first-person perspective. In light of the emerging results, it will be interesting to implement the paradigm by administering questionnaires about the sense of agency and the illusion of control over the task to relate them to our current hypotheses on the peculiar increase of the P2/3a registered in the Auditory Motor Condition.

4.2 Limitations

The study presents some limitations that should be deemed when interpreting the results; different criticalities appear at various stages of the research and will be spotlighted throughout this paragraph.

Limitations related to data collection. First, the data were collected from a sample of undergraduate students, with most of them having a psychology studies background that could have influenced their naive approach to the paradigm. In addition, some participants performed similar studies in the past and might have gained familiarity with the procedure. Also, the language could have been a problem for the non-native German participants recruited from the international network of the Heinrich Heine University; indeed, being exposed to the study instruction in German might have influenced their performance. As evidence of this, some participants showed that they did not fully understand the instructions; a couple of them needed to repeat the experimental procedure as they had unclear when or how to press the button (lately, these two participants were excluded from the analyses because their responses were too slow to be analyzed). Despite the attempts to overcome the misunderstandings, other miscomprehensions might have appeared and could not have been controlled by the researcher since the subject completed the paradigm in an isolated room and could not ask for clarifications. Fatigue and boredom might also have played a role in their performance; during the first part of the experimental session, participants were prepared for the EEG registration, increasing their tiredness; also, the likelihood of being distracted was very high: throughout the experimental paradigm, the button press was the only action to perform, not even required in every block. During the AVC and the AOC, the subjects were in fact only passive observers. Nonetheless, giving a final reward could have motivated them and might have helped them to maintain a higher level of attention.

Moreover, the sample was not representative of the population, composed of a group of students from eighteen to twenty-five years old; for the results to be generalized, a more representative sample might be necessary. Additionally, we did not perform any power analysis during the study preparation. We started the data collection aiming to test at least twenty participants, taking as reference the study of SanMiguel et al. (2013) and considering that most of the studies using neuroimaging techniques include twenty to twenty-five subjects. Despite this, a power analysis could have helped to check for the goodness of the procedure. Additionally, the gender could have been better balanced within the sample, comprising fourteen females and four males

if we consider the biological sex, while fourteen females, three males, and one other for the social sex. Post-hoc analysis and further studies might consider the potential effect of gender on the results. Another plausible problem concerns the presence of some research biases, namely all the factors that can lead to misleading conclusions during the research process (Šimundić, 2013). In the current study, data collection was always in charge of the same person, which is optimal to increase the internal validity of the study; however, tiredness should have played a role, especially during the montage and the setup of the experimental procedure. In addition, dealing with an experimenter who was a non-native German speaker might have exerted an influence since most participants were from Germany: being forced to speak in English might also have increased their level of anxiety.

Limitations related to data registration. Regarding the EEG signal, low-frequency ERPs do not reflect the complete neuronal response for prediction errors. Despite the limits related to invasiveness, high-frequency functional recordings such as Electrocorticography (ECoG) prove to be a reliable alternative to the classic scalp-electrode recordings. They allow the comparison of a restricted pattern of activity from specific brain regions, which results in a higher spatial resolution (Dürschmid et al., 2016). ECoG consists of the direct recording of cortical activity with subdural electrodes, and it can better suit the study of PE in the auditory domain: it underlines an increase in the γ high-frequency band (80-150 Hz) for deviant signals in frontal and temporal areas (Dürschmid et al., 2016). Changes in the γ band manifest earlier than ERPs variations. as a result, ECoG detects frequency changes with high temporal resolution. Conversely, the classic scalp EEG recording depicts a general picture of the cortical activity during PE transmission, considering whole-brain electrophysiological changes (Luck, 2014). To date, high-frequency functional recordings have been used to study auditory PE signals for deviants, but not for sound predicted by self-induced cues (Dürschmid et al., 2016). This technique should be a valuable means to expand the current knowledge about specific signal variations related to prediction error changes. Last, participants were observing the monitor for the whole duration of the study without being required to do anything else; for this reason, we must be sure that their gazes pointed toward the screen and the presented stimuli. The use of eye tracking can help reduce the variance linked to this process, especially during the observation of visual cues, as previously done by past studies on predictive processing conducted in this laboratory.

Limitations related to statistical analyses. Other considerations about the limits pertain to the statistical analyses. First, each experimental block was presented only once during the paradigm to allow for a smoother calculation of prediction error indexes at the expense of the internal study validity and in favor of a random variance increase. Future studies should present each block more than once to guarantee a higher statistical power of the results. A Second problem originates from the motor correction of the AMC signal, performed without meticulously following the procedure described by Horvath (2015). In his article about the Contingent paradigm methodology, the author states the necessity to verify the presence of any difference in the button press timing between the AMC and the MOC: indeed, if the timing is statistically different between the two, they might underpin unrelated elaboration processes (Horváth, 2015). Despite this, this type of motor correction has been performed in many studies using the button press paradigm, which makes the procedure methodologically safe. The button press timing verification is a hint to increase the internal validity.

Moreover, the mixed linear models only include the factor Subjects within the random effects, which allows for controlling the random variance from individual differences. Originally, the electrodes were also embraced within the random effects but then discarded because the models did not reach convergence by including them. For a better evaluation of the results, future studies should consider this factor again by analyzing data from a broader sample of participants so that the convergence of the model is easier to achieve. For these reasons, the performance of a power analysis before data collection helps to have a clearer idea of the number of participants for setting a statistical model that fits all the conditions. Last, the electrophysiological analyses focus on three frontocentral electrodes only. Whole-brain analysis of the electrophysiological data might be the best solution to underpin the entire N1 and P2 intercourse, trying to avoid biases from the preselection of some components.

4.3 Conclusions

This study shows a significant increase in P2 amplitude during auditory events processing that correlates positively with the magnitude of prediction error indexes. These results may reflect the feeling of having poor control over the predictability of the task and the urge to update the internal cognitive model to overcome this uncertainty. The P2 increase in the AMC might also be related to the focus of attentional resources on the auditory stimulus and the activation of high-level computational processes. Future studies increasing the presentation rate of the sound so that it appears in more than 50% of trials may help to understand whether P2 amplitude decreases when the target stimulus is more clearly associated with the button press. In addition, varying the tone rate may help clarify the relationship between visual cues and P2 changes during the AVC, which remains unclear. Also, the introduction of questionnaires on the perceived sense of control may further complete the paradigm and confirm our hypothesis about the cognitive computations associated with P2 changes. All these precautions, together with the adjustments reported in the Limitations section, will allow us to obtain a clear picture of the results and to fit auditory predictive processing into the most appropriate theoretical framework. The current results support the theories that claim a specific predictive processing pathway for motor computations, as visual cues do not appear to be involved in predictive processing.

One major weakness of this project, in my opinion, is the lack of behavioral indices that correlate with the electrophysiological data. This step has also been neglected in other studies about auditory predictive processing but is essential for increasing the internal validity of the paradigm and the clarity of the cognitive processes studied. Indeed, prediction error indices were calculated using a mathematical formula and spontaneously associated with EEG data. The association of a behavioral index reflecting the ongoing computation of predictions would increase the strength of the paradigm. Nonetheless, this pilot study aimed to test whether the modified version of the classical button press paradigm could be suitable to investigate predictive processing in the auditory domain by implementing a trial-by-trial analysis. Because of its exploratory goal, the project was not even submitted to the ethics committee; regardless of any reservations, the resulting considerations are a valuable starting point for future investigations of predictive processing in the auditory domain.

Appendix

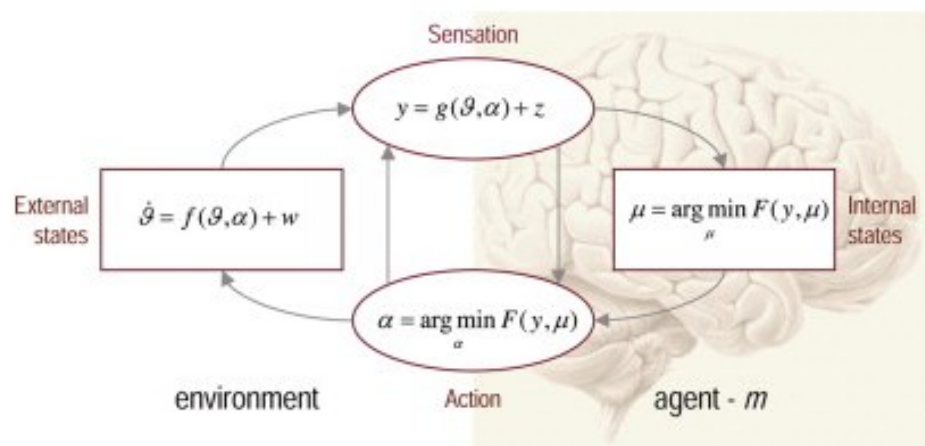
5.1 Electroencephalography: an overview

Electroencephalography (EEG) is a non-invasive method for recording the spontaneous fluctuations of cortical electrical potentials detected through surface electrodes. It is commonly used for clinical and research purposes, thanks to its reliability and non-invasiveness. More precisely, EEG depicts the summation of the postsynaptic potentials at the level of the extracellular fluid; thus, it does not reflect the direct brain activity represented by the action potentials, but it is just an indirect measure of it. The raw electrical signal undergoes a Fourier transform that allows distinguishing waves of different frequencies. Compared to the other brain-imaging techniques, EEG has an optimal temporal but poor spatial resolution (Burle et al., 2015). It captures the brain dynamics with a temporal precision in the order of milliseconds; however, inferences about the signal origins cannot be possible since there is no direct relationship between the signal from a given electrode and the activity of the area underneath the electrode. EEG is not suitable for detecting the dynamics from deep layers, which undergo dispersion while reaching the scalp. Thus, it is only appropriate for registering the electric signal from surface areas. For the registration to be efficient, the electrodes should be placed perpendicularly to the scalp; moreover, the spatial resolution and quality of the signal increase with the number of electrodes. The current study uses a 28-electrodes montage, but more sophisticated caps contain up to 64 electrodes. Passive ring electrodes are used, transmitting the voltage fluctuations to the amplifier through a conductive wire. Electrodes are made from good conductors such as Ag, AgCl, or gold to attain a low impedance. As said before, the montage follows the International 10-20 system, which implies the placement of electrodes at proportional distances along two hypothetical lines connecting four reference points: the sagittal line from inion to nasion and the transversal line that links the two auricular depressions. The montage is monopolar: the signal is the difference between the ground electrode, placed in AFz, with each active electrode; in the case of bipolar montages, it reflects the difference between pairs of active electrodes. It is referenced to the mastoids, whose electric signal is registered by placing two electrodes behind the ears. This reference is widely used, but it can have some drawbacks when the subject is tense and the mastoids contract, corrupting the signal. Lastly, registering the electrooculogram with the brain signal is essential since eye movement constitutes an artifact. For this reason, it is better to detect the artefactual signal from eye movements and remove it after the acquisition than ask the subject not to move the eyes. The

latter solution allows getting a clean signal but interferes with the spontaneousness of the task and adds cognitive effort to the participant.

5.2 The free-energy principle

The free-energy principle postulated by Friston et al. proposes to explain adaptation processes, extending to the biological systems and networks unrelated to the human brain. It assumes that all the adaptive changes related to a system lead to a minimization of its overall energy; when implementing the principle at a brain level, the amount of free energy relies on the computations linking sensory data and brain states. This definition is similar to prediction errors; following specific assumptions, the two concepts can be considered synonyms. In the first mathematical implementation of the principle, Friston relies on two tenets: (I) agents try to resist the natural tendency to entropy by minimizing the level of energy inside a system; (II) in reaching this goal, they act upon the environment to avoid surprises. The structure of the formula represents the computations leading to the system adaptation, while the factors inside it are the stimuli under elaboration. Figure 5.1, from Friston et al. (2009), shows the basic mathematics behind the free-energy principle. The three main factors contributing to the final adaptation are actions, sensations, and external states; they can assume different values, contributing to changes in the free energy state (F). The final value of F is the result of the difference between the total energy and the amount of entropy within the system.



$$F = \text{Energy} - \text{Entropy} = -\langle \ln p(y, \theta | m) \rangle_q + \langle \ln q(\theta) \rangle_q$$

Figure 5.1 Mathematical formulation of the free-energy principle by Friston et al. (2009).

Within this first formulation, the final brain state is the μ factor; the experience of the environment, parametrized by the factor ϑ , depends on the action made to explore it, indicated by α . The formula is a good representation of the interactions between the brain states and the environmental stimuli, but no predictions are considered. Additionally, it is a good proxy of classic internal forward models (Miall & Wolpert, 1996), in which actions are the only means to reduce the entropy and guide the perceptual experience. In the following conceptualizations of the mathematical model, Friston introduces an additional core concept, according to which individuals trust Bayesian inferences (Knill & Pouget, 2004) to guide their interactions with the environment and make predictions about unknown experiences. The tenet needs to be updated

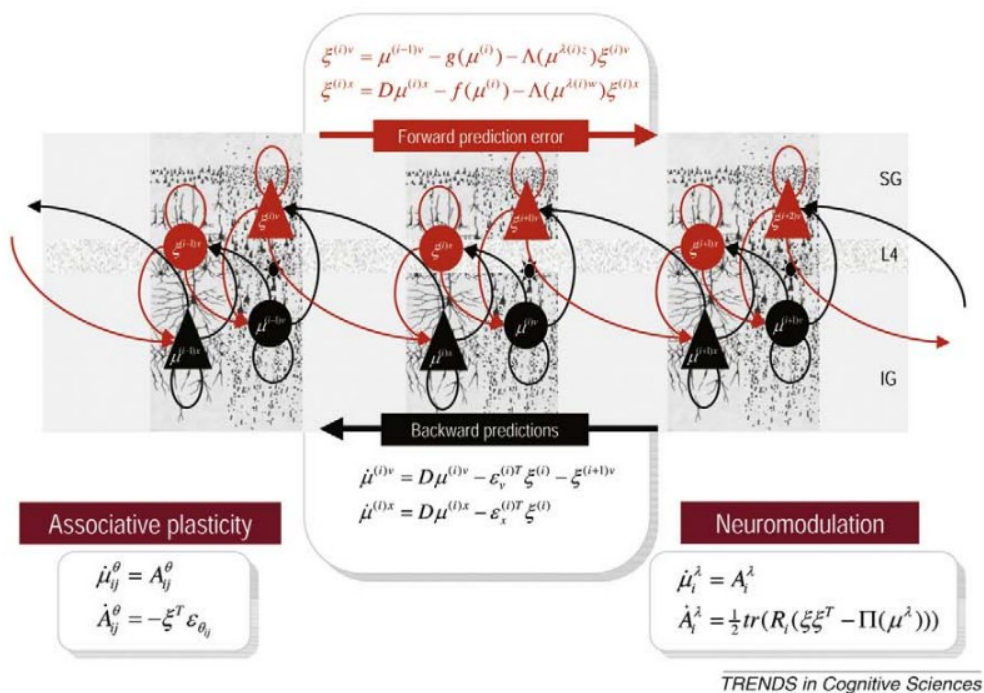


Figure 5.2 Second formulation of the free energy principle and its neurobiological application. The formula is updated considering the presence of Bayesian inferences. The image is taken by Friston et al. (2009).

to the previous formula, considering the constant attempts of the brain to get the likelihood of obtaining some data given their causes; a Bayesian equation entailing the factors abovementioned adapts to these needs. The final formula, reported in Figure 5.2 by Friston (2009), considers this new adaptation and implements the mathematics behind it at a neurobiological level, considering its application for different neuronal populations. Predictions (ξ) are processed inside each layer and are the result of the interaction between prediction errors from the lower level (μ^{i-1}) and the computations coming from the current hierarchical level. The factors g and f represent the top-down and lateral connections inside each layer, respectively.

Additionally, π accounts for the precision weighting, associated with sensory attenuation and attentional mechanisms (Chennu et al., 2013; Friston, 2018); the factor θ is the quantity of random variance contributing to the final result. The formula provides a good representation of the theoretical underpinnings behind Predictive Coding (Rao & Ballard, 1999).

Future mathematical models should integrate the computation occurring within the same anatomical structures, such as learning, allowing an embedded view of all the brain computations contributing to the creation of the experience (Friston, 2018).

5.3 The Edinburgh Handedness Inventory

The Edinburgh Handedness Inventory (EHI) was developed by Oldfield (1971); it assesses the handedness of a subject in a set of activities of daily living. The original version of the Inventory includes the 20 following items. The rater can assign two crosses for the preferential side, one cross for the side if there is no preference, or no crosses if there is no familiarity with the action.

1. Writing
2. Drawing
3. Throwing
4. Scissors
5. Comb
6. Toothbrush
7. Knife (without fork)
8. Spoon
9. Hammer
10. Screwdriver
11. Tennis racket
12. Knife with fork
13. Cricket bat
14. Golf club
15. Broom
16. Rake
17. Striking a match
18. Opening a box (lid)
19. Dealing cards
20. Threading a needle

From the above, items 1, 2, 3, 4, 6, 7, 8, 15, 17, and 18 made it into the 10-item version of the EHI that is the one administered in the pilot study. Additionally, two items are added to the original: the preferred side for kicking and a question about the dominant eye. The final score, called Laterality Quotient, is calculated by using the formula: $(R-L)/(R+L) \times 100$, with R and L referring to the sum of the crosses assigned for the right (R) and the left (L) sides.

5.4 R scripts for the statistical analyses

The following R scripts, developed on R Studio, report the statistical analyses performed on the electrophysiological data considering the components of interest, namely N1 and P2. The third R script was applied to P2 data for exploratory purposes only and is not mentioned in the preceding chapters. Only the scripts applied to P2 data are reported; two analogous scripts using a different data set and naming the variables as n1m and n1m_noEXT are used for the analysis of N1.

1. Sound type (AMC, AVC, AOC)

The following script investigates the effect of Sound type, considering the three levels within this variable (AMC, AVC, AOC). Here, PE is not considered an independent variable.

```
# Read Data
p2mdat1 <- read_excel ("single_trial_164_268_pos_40_cor_trials_2_5SD_data_clean.xlsx", sheet = 1, col_names = TRUE, col_types = NULL, na = "", skip = 0)
p2mdat1 <- (data.frame(p2mdat1))
p2mdat1 <- subset (p2mdat1, subject != "AO12TO")

#Data Recoding
p2mdat1 <- within (p2mdat1, {
  #Tdcs
  sound_type.f <- factor(condition, levels = c("EXT","ACT","VIS"))
  EXT.sound_type.f <- sound_type.f
  ACT.sound_type.f <- sound_type.f
  VIS.sound_type.f <- sound_type.f
  contrasts(sound_type.f) = contr.treatment(3, base=1) - matrix(rep(1/3, 6), ncol=2)
  contrasts (EXT.sound_type.f) = contr.treatment(3, base=1)
  contrasts (ACT.sound_type.f) = contr.treatment(3, base=2)
  contrasts (VIS.sound_type.f) = contr.treatment(3, base=3)
  # Prediction Error
  c.PE <- PE - mean (PE, na.rm=T)
  low.PE <- c.PE + sd (c.PE, na.rm=T)
  high.PE <- c.PE - sd (c.PE, na.rm=T)
})
count (p2mdat1, c ("subject"))

# Three-level analysis
p2m = lmer (value ~ sound_type.f + (1 + sound_type.f|subject), REML=TRUE, data = p2mdat1, control=(lmerControl(optimizer ="bobyqa"))

anova(p2m)
summary(p2m)
ggeffect(p2m, terms = c("sound_type.f"), ci.lvl = 0.95)
```

2. Sound type (AMC, AVC) x PE

The following R script examines the effect of Sound type, PE, and their interaction. AOC is removed from the levels of the independent variable Sound type.

```
# Read Data
p2m_noEXTdata1 <- read_excel ("single_trial_164_268_pos_40_cor_trials_2_5SD_data_clean.xlsx", sheet = 1, col_names = TRUE, col_types = NULL, na = "", skip = 0)
p2m_noEXTdata1 <- (data.frame(p2m_noEXTdata1))
p2m_noEXTdata1 <- subset (p2m_noEXTdata1, condition != 'EXT')
p2m_noEXTdata1 <- subset (p2m_noEXTdata1, subject != "AO12TO")

#Data Recoding
p2m_noEXTdata1 <- within (p2m_noEXTdata1, {
# Tdcs
  sound_type.f <- factor(condition, levels = c("VIS","ACT"))
  ACT.sound_type.f <- sound_type.f
  VIS.sound_type.f <- sound_type.f
  contrasts (sound_type.f) = contr.treatment(2, base=1) - matrix(rep(1/2, 2), ncol=1)
  contrasts (ACT.sound_type.f) = contr.treatment(2, base=2)
  contrasts (VIS.sound_type.f) = contr.treatment(2, base=1)
# Prediction Error
  c.PE <- PE - mean (PE, na.rm=T)
  low.PE <- c.PE + sd (c.PE, na.rm=T)
  high.PE <- c.PE - sd (c.PE, na.rm=T)
})
Count (p2m_noEXTdata1, c("subject"))

# Three level analysis
p2m_noEXT = lmer(value ~ c.PE*sound_type.f + (1 + sound_type.f * c.PE|subject), REML=TRUE, data = p2m_noEXTdata1, control=(lmerControl(optimizer = "bobyqa")))

anova(p2m_noEXT)
summary(p2m_noEXT)
ggeffect(p2m_noEXT, terms = c("sound_type.f","c.PE [-0.25, 0.25]"), ci.lvl = 0.95)
emmeans(p2m_noEXT, pairwise~ c.PE + sound_type.f, pbkrtest.limit = 5041)

# Single level analysis
p2m_noEXT_ACT = lmer(value ~ c.PE*ACT.sound_type.f + (1 + sound_type.f * c.PE|subject), REML=TRUE, data = p2m_noEXTdata1, control=(lmerControl(optimizer = "bobyqa")))
p2m_noEXT_VIS = lmer(value ~ c.PE*VIS.sound_type.f + (1 + sound_type.f * c.PE|subject), REML=TRUE, data = p2m_noEXTdata1, control=(lmerControl(optimizer = "bobyqa")))

summary(p2m_noEXT_ACT)
summary(p2m_noEXT_VIS)
anova(p2m_noEXT_ACT)
anova(p2m_noEXT_VIS)
```

```

#Analysis with high and low PE values
p2m_noEXT_lowPE = lmer(value ~ low.PE*sound_type.f + (1 + sound_type.f *
c.PE|subject), REML=TRUE, data = p2m_noEXTdata, control=(lmerControl(opti-
mizer ="bobyqa"))
p2m_noEXT_highPE = lmer(value ~ high.PE*sound_type.f + (1 + sound_type.f *
c.PE|subject), REML=TRUE, data = p2m_noEXTdata, control=(lmerControl(opti-
mizer ="bobyqa"))

summary(p2m_noEXT_lowPE)
summary(p2m_noEXT_highPE)
anova(p2m_noEXT_highPE)
anova(p2m_noEXT_lowPE)

```

3. Sound type (AMC, AVC, AOC) x PE

The following R script is not used in the final analyses and thus not discussed in the previous chapters. It was performed on P2 data only (due to the high significance of the independent variables found in the previous analyses) to have a clear idea of the effects of the independent variables and the contrasts between the Sound type levels when considering AOC, which acts as the control condition.

```

# Read Data
p2mdat1 <- read_excel ("single_trial_164_268_pos_40_cor_tri-
als_2_5SD_data_clean.xlsx", sheet = 1, col_names = TRUE, col_types = NULL, na =
"", skip = 0)
p2mdat1 <- (data.frame(p2mdat1))
p2mdat1 <- subset (p2mdat1, subject != "AO12TO")

#Data recording
p2mdat1 <- within (p2mdat1, {
# tdc
sound_type.f <- factor(condition, levels = c("EXT","ACT","VIS"))
EXT.sound_type.f <- sound_type.f
ACT.sound_type.f <- sound_type.f
VIS.sound_type.f <- sound_type.f
contrasts (sound_type.f) = contr.treatment(3, base=1) - matrix(rep(1/3, 6), ncol=2)
contrasts (EXT.sound_type.f) = contr.treatment(3, base=1)
contrasts (ACT.sound_type.f) = contr.treatment(3, base=2)
contrasts (VIS.sound_type.f) = contr.treatment(3, base=3)
# Prediction Error
c.PE <- PE - mean (PE, na.rm=T)
low.PE <- c.PE + sd(c.PE, na.rm=T)
high.PE <- c.PE - sd(c.PE, na.rm=T)
})
count (p2mdat1,c("subject"))

```

```

# Three level analysis
p2m = lmer (value ~ c.PE*sound_type.f + (1 + sound_type.f * c.PE|subject),
REML=TRUE, data = p2mdatal, control=(lmerControl(optimizer ="bobyqa")))

anova(p2m)
summary(p2m)

# Single level analysis
p2m_EXT = lmer (value ~ c.PE*EXT.sound_type.f + (1 + sound_type.f + c.PE|sub-
ject), REML=TRUE, data = p2mdatal, control=(lmerControl(optimizer ="bobyqa")))
p2m_ACT = lmer (value ~ c.PE*ACT.sound_type.f + (1 + sound_type.f + c.PE|sub-
ject), REML=TRUE, data = p2mdatal, control=(lmerControl(optimizer ="bobyqa")))
p2m_VIS = lmer (value ~ c.PE*VIS.sound_type.f + (1 + sound_type.f + c.PE|subject),
REML=TRUE, data = p2mdatal, control=(lmerControl(optimizer ="bobyqa")))

summary(p2m_EXT)
summary(p2m_ACT)
summary(p2m_VIS)

```

5.5 Mixed Linear Models performed on R studio

Table 4. reports the formulas used as inputs of the *lmer* function in R studio. Sound type and PE represent the fixed effects, while the Subject is the random effect. The Mixed Linear Models are fit according to the Restricted Maximum Likelihood (REML) criteria, while the t-tests use Satterthwaite's method. The BOBYQA optimizer optimizes the MLM computations by quadratic approximation.

Table 4. Formulas designed for each Mixed Linear Model.

Models	
N1 analysis	MLM for Sound type (AOC, AVC,AMC): value ~ sound_type.f + (1 + sound_type.f subject)
	MLM for Sound type (AVC, AMC) x PE value ~ c.PE * sound_type.f + (1 + sound_type.f * c.PE subject)
P2 analysis	MLM for Sound type (AOC, AVC,AMC) value ~ sound_type.f + (1 + sound_type.f subject)
	MLM for Sound type (AVC, AMC) x PE value ~ c.PE * sound_type.f + (1 + sound_type.f * c.PE subject)
	value ~ c.PE*ACT.sound_type.f + (1 + sound_type.f * c.PE subject)
	value ~ c.PE*VIS.sound_type.f + (1 + sound_type.f * c.PE subject)
	value ~ low.PE*sound_type.f + (1 + sound_type.f * c.PE subject)
value ~ high.PE*sound_type.f + (1 + sound_type.f * c.PE subject)	

References

6.1 Bibliography

- Abolafia, J. M., Vergara, R., Arnold, M., Reig, R., & Sanchez-Vives, M. V. (2011). Cortical auditory adaptation in the awake rat and the role of potassium currents. *Cerebral Cortex*.
<https://doi.org/10.1093/cercor/bhq163>
- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9, 357–381. <https://doi.org/10.1146/annurev.ne.09.030186.002041>
- Alho, K., Rinne, T., Herron, T. J., & Woods, D. L. (2014). Stimulus-dependent activations and attention-related modulations in the auditory cortex: A meta-analysis of fMRI studies. *Hearing Research*, 307, 29–41. <https://doi.org/10.1016/j.heares.2013.08.001>
- 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>
- Andreasen, N. C. (1997). The role of the thalamus in schizophrenia. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 42(1), 27–33.
<https://doi.org/10.1177/070674379704200104>
- 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>
- Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398. <https://doi.org/10.1016/j.tics.2012.05.003>

- Arnal, L. H., Wyart, V., & Giraud, A.-L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, *14*(6), 797–801.
<https://doi.org/10.1038/nn.2810>
- Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*, *48*(9), 1276–1283.
<https://doi.org/10.1111/j.1469-8986.2011.01196.x>
- Bansal, S., Ford, J. M., & Spering, M. (2018). The function and failure of sensory predictions. *Annals of the New York Academy of Sciences*, *1426*(1), 199–220.
<https://doi.org/10.1111/nyas.13686>
- Barbas, H., & Rempel-Clower, N. (1997). Cortical structure predicts the pattern of corticocortical connections. *Cerebral Cortex*, *7*(7), 635–646. <https://doi.org/10.1093/cercor/7.7.635>
- 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: Official Journal of the International Organization of Psychophysiology*, *70*(2), 137–143.
<https://doi.org/10.1016/j.ijpsycho.2008.06.005>
- 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>
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, *76*(4), 695–711.
<https://doi.org/10.1016/j.neuron.2012.10.038>

- Bays, P. M., Flanagan, J. R., & Wolpert, D. M. (2006). Attenuation of Self-Generated Tactile Sensations Is Predictive, not Postdictive. *PLoS Biology*, 4(2), e28.
<https://doi.org/10.1371/journal.pbio.0040028>
- Behroozmand, R., & Larson, C. R. (2011a). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, 12(54).
<https://doi.org/10.1186/1471-2202-12-54>
- Behroozmand, R., & Larson, C. R. (2011b). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, 12(1), 54.
<https://doi.org/10.1186/1471-2202-12-54>
- Bellebaum, C., Ghio, M., Wollmer, M., Weismüller, 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>
- Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, 83(2), 120–131. <https://doi.org/10.1016/j.ijpsycho.2011.08.003>
- Berkeley, G. (1709). *An essay towards a new theory of vision*. Dublin: printed by Aaron Rhames, for Jeremy Pepyat, 1709. <http://name.umdl.umich.edu/004848514.0001.000>
- Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-Temporal Prediction Modulates the Perception of Self-Produced Stimuli. *Journal of Cognitive Neuroscience*, 11(5), 551–559.
<https://doi.org/10.1162/089892999563607>
- Bonneh, Y. S., Donner, T. H., Sagi, D., Fried, M., Cooperman, A., Heeger, D. J., & Arieli, A. (2010). Motion-induced blindness and microsaccades: Cause and effect. *Journal of Vision*, 10(14), 22. <https://doi.org/10.1167/10.14.22>
- Burle, B., Spieser, L., Roger, C., Casini, L., Hasbroucq, T., & Vidal, F. (2015). Spatial and temporal resolutions of EEG: Is it really black and white? A scalp current density view.

International Journal of Psychophysiology, 97(3), 210–220.

<https://doi.org/10.1016/j.ijpsycho.2015.05.004>

Buschman, T. J., & Miller, E. K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science*, 315(5820), 1860–1862.

<https://doi.org/10.1126/science.1138071>

Cao, L., & Gross, J. (2015). Cultural Differences in Perceiving Sounds Generated by Others: Self Matters. *Frontiers in Psychology*, 6.

<https://www.frontiersin.org/article/10.3389/fpsyg.2015.01865>

Cao, L., Thut, G., & Gross, J. (2017). The role of brain oscillations in predicting self-generated sounds. *NeuroImage*, 147, 895–903. <https://doi.org/10.1016/j.neuroimage.2016.11.001>

Carbajal, G. V., & Malmierca, M. S. (2018). The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection. *Trends in Hearing*, 22, 2331216518784822. <https://doi.org/10.1177/2331216518784822>

Catani, Marco, and Michel Thiebaut de Schotten, *Atlas of Human Brain*

Connections (Oxford, 2012; online edn, Oxford Academic, 1 Nov.

2012), <https://doi.org/10.1093/med/9780199541164.001.0001>

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., Owen, A. M., & Bekinschtein, T. A. (2013). Expectation and Attention in Hierarchical Auditory Prediction.

Journal of Neuroscience, 33(27), 11194–11205. <https://doi.org/10.1523/JNEUROSCI.0114-13.2013>

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. MIT Press.

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, 9(8), 587–600. <https://doi.org/10.1038/nrn2457>

- Darriba, Á., Hsu, Y.-F., Van Ommen, S., & Waszak, F. (2021). Intention-based and sensory-based predictions. *Scientific Reports*, *11*(1), 19899. <https://doi.org/10.1038/s41598-021-99445-z>
- Davis, H., & Zerlin, S. (1966). Acoustic Relations of the Human Vertex Potential. *The Journal of the Acoustical Society of America*, *39*(1), 109–116. <https://doi.org/10.1121/1.1909858>
- Den Ouden, H., Kok, P., & De Lange, F. (2012). How Prediction Errors Shape Perception, Attention, and Motivation. *Frontiers in Psychology*, *3*.
<https://www.frontiersin.org/article/10.3389/fpsyg.2012.00548>
- Dogge, M., Custers, R., & Aarts, H. (2019a). 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., Custers, R., Gayet, S., Hoijtink, H., & Aarts, H. (2019b). Perception of action-outcomes is shaped by life-long and contextual expectations. *Scientific Reports*, *9*(1), 5225.
<https://doi.org/10.1038/s41598-019-41090-8>
- Dogge, M., Hofman, D., Custers, R., & Aarts, H. (2019c). 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>
- DSM-5 American Psychiatric Association, Manuale diagnostico e statistico dei disturbi mentali (2013). Raffaello Cortina Editore.
- Duque, D., Ayala, Y. A., & Malmierca, M. S. (2015). Deviance detection in auditory subcortical structures: What can we learn from neurochemistry and neural connectivity? *Cell and Tissue Research*, *361*(1), 215–232. <https://doi.org/10.1007/s00441-015-2134-7>
- Dürschmid, S., Edwards, E., Reichert, C., Dewar, C., Hinrichs, H., Heinze, H.-J., Kirsch, H. E., Dalal, S. S., Deouell, L. Y., & Knight, R. T. (2016). Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *Proceedings of the National Academy*

of Sciences of the United States of America, 113(24), 6755–6760.

<https://doi.org/10.1073/pnas.1525030113>

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., & Giraud, A.-L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature Communications*, 5(1), 4694. <https://doi.org/10.1038/ncomms5694>

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>

Frey, S., Campbell, J. S. W., Pike, G. B., & Petrides, M. (2008). Dissociating the Human Language Pathways with High Angular Resolution Diffusion Fiber Tractography. *Journal of Neuroscience*, 28(45). <https://doi.org/10.1523/jneurosci.2388-08.2008>

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, 16(9), 1325–1352. <https://doi.org/10.1016/j.neunet.2003.06.005>

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301. <https://doi.org/10.1016/j.tics.2009.04.005>

Friston, K. (2018). Does predictive coding have a future? *Nature Neuroscience*, 21(8), 1019–1021. <https://doi.org/10.1038/s41593-018-0200-7>

Friston, K. J., Shiner, T., FitzGerald, T., Galea, J. M., Adams, R., Brown, H., Dolan, R. J., Moran, R., Stephan, K. E., & Bestmann, S. (2012). Dopamine, Affordance and Active Inference. *PLOS Computational Biology*, 8(1), e1002327. <https://doi.org/10.1371/journal.pcbi.1002327>

Gentsch, A., & Schütz-Bosbach, S. (2011). I Did It: Unconscious Expectation of Sensory Consequences Modulates the Experience of Self-agency and Its Functional Signature. *Journal of Cognitive Neuroscience*, 23(12), 3817–3828. https://doi.org/10.1162/jocn_a_00012

- 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., 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>
- Haggard, P., & Whitford, B. (2004). Supplementary motor area provides an efferent signal for sensory suppression. *Cognitive Brain Research*, 19(1), 52–58. <https://doi.org/10.1016/j.cogbrainres.2003.10.018>
- Halgren, E., Marinkovic, K., & Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and Clinical Neurophysiology*, 106(2), 156–164. [https://doi.org/10.1016/S0013-4694\(97\)00119-3](https://doi.org/10.1016/S0013-4694(97)00119-3)
- 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>
- Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience*, 389, 54–73. <https://doi.org/10.1016/j.neuroscience.2017.07.061>
- Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, 42(2), 180–190. <https://doi.org/10.1111/j.1469-8986.2005.00272.x>

- Henson, R. N. A., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*(3), 263–270.
[https://doi.org/10.1016/S0028-3932\(02\)00159-8](https://doi.org/10.1016/S0028-3932(02)00159-8)
- 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>
- 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>
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719.
<https://doi.org/10.1016/j.tins.2004.10.007>
- Knolle, F., Schröger, E., & Kotz, S. A. (2013a). Prediction errors in self- and externally-generated deviants. *Biological Psychology*, *92*(2), 410–416.
<https://doi.org/10.1016/j.biopsycho.2012.11.017>
- Knolle, F., Schröger, E., & Kotz, S. A. (2013b). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*(9), 2449–2461. <https://doi.org/10.1016/j.cortex.2012.12.012>
- 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>
- Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–579.
[https://doi.org/10.1016/S0166-2236\(00\)01657-X](https://doi.org/10.1016/S0166-2236(00)01657-X)
- 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>

- Lenartowicz, A., Escobedo-Quiroz, R., & Cohen, J. D. (2010). Updating of context in working memory: An event-related potential study. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 298–315. <https://doi.org/10.3758/CABN.10.2.298>
- Lindner, A., Thier, P., Kircher, T., Haarmeier, T., & Leube, D. (2005). Disorders of Agency in Schizophrenia Correlate with an Inability to Compensate for the Sensory Consequences of Actions. *Current Biology*. <https://doi.org/10.1016/j.cub.2005.05.049>
- Luck, S. J. (2005). *An Introduction to the Event-related Potential Technique*. MIT Press.
- Malpas, J., Donald Davidson (Winter 2012 Edition), The Stanford Encyclopedia of Philosophy (Winter 2012 Edition), Edward N. Zalta.
- Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed Responses to Self-triggered Sounds in the Human Auditory Cortex. *Cerebral Cortex*, *15*(3), 299–302. <https://doi.org/10.1093/cercor/bhh131>
- Martikainen, M., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*, 299–302.
- Maunsell, J. H. R., & Essen, D. C. V. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, *3*(12), 2563–2586. <https://doi.org/10.1523/jneurosci.03-12-02563.1983>
- 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>
- Miall, R. C., & Wolpert, D. M. (1996). Forward Models for Physiological Motor Control. *Neural Networks*, *9*(8), 1265–1279. [https://doi.org/10.1016/S0893-6080\(96\)00035-4](https://doi.org/10.1016/S0893-6080(96)00035-4)
- Minati, L., Salvatoni, L., Rosazza, C., Pietrocini, E., Visani, E., Panzica, F., Scaioli, V., Avanzini, G., & Franceschetti, S. (2010). Event-related potential (ERP) markers of melodic processing:

- The N2 component is modulated by structural complexity, not by melodic ‘meaningfulness’.
Brain Research Bulletin, 83(1), 23–28. <https://doi.org/10.1016/j.brainresbull.2010.06.011>
- Moore, J. W., Lagnado, D., Deal, D. C., & Haggard, P. (2009). Feelings of control: Contingency determines experience of action. *Cognition*, 110(2), 279–283.
<https://doi.org/10.1016/j.cognition.2008.11.006>
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics*, 66(3), 241–251. <https://doi.org/10.1007/BF00198477>
- Näätänen, R., Astikainen, P., Ruusuvirta, T., & Huotilainen, M. (2010). Automatic auditory intelligence: An expression of the sensory–cognitive core of cognitive processes. *Brain Research Reviews*, 64(1), 123–136. <https://doi.org/10.1016/j.brainresrev.2010.03.001>
- Neisser, U. (1967). *Cognitive psychology*. New York, Appleton-Century-Crofts.
<http://archive.org/details/cognitivepsychol00neis>
- 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>
- Numminen, J., Salmelin, R., & Hari, R. (1999). Subject’s own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*, 265(2), 119–122.
[https://doi.org/10.1016/S0304-3940\(99\)00218-9](https://doi.org/10.1016/S0304-3940(99)00218-9)
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *International Journal of Psychophysiology*, 88(2), 109–123. <https://doi.org/10.1016/j.ijpsycho.2013.03.015>
- Parras, G. G., Nieto-Diego, J., Carbajal, G. V., Valdés-Baizabal, C., Escera, C., & Malmierca, M. S. (2017). Neurons along the auditory pathway exhibit a hierarchical organization of

prediction error. *Nature Communications*, 8(1), 2148. <https://doi.org/10.1038/s41467-017-02038-6>

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>

Pinheiro, A. P., Schwartz, 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>

Raij, T., McEvoy, L., Mäkelä, J. P., & Hari, R. (1997). Human auditory cortex is activated by omissions of auditory stimuli. *Brain Research*, 745(1), 134–143. [https://doi.org/10.1016/S0006-8993\(96\)01140-7](https://doi.org/10.1016/S0006-8993(96)01140-7)

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>

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, 4059. <https://doi.org/10.1038/ncomms5059>

Reznik, D., & Mukamel, R. (2019). Motor output, neural states and auditory perception. *Neuroscience & Biobehavioral Reviews*, 96, 116–126. <https://doi.org/10.1016/j.neubiorev.2018.10.021>

Reznik, D., Ossmy, O., & Mukamel, R. (2015). Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *Journal of Neuroscience*, 35(5), 2173–2180. <https://doi.org/10.1523/JNEUROSCI.3723-14.2015>

- Reznik, D., Simon, S., & Mukamel, R. (2018). Predicted sensory consequences of voluntary actions modulate amplitude of preceding readiness potentials. *Neuropsychologia*, *119*, 302–307. <https://doi.org/10.1016/j.neuropsychologia.2018.08.028>
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. <https://doi.org/10.1038/14819>
- SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing Silences: Human Auditory Processing Relies on Preactivation of Sound-Specific Brain Activity Patterns. *Journal of Neuroscience*, *33*(20), 8633–8639. <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>
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, *105*(46), 18035–18040. <https://doi.org/10.1073/pnas.0805234105>
- Schafer, E. W. P. (1982). Neural Adaptability: A Biological Determinant of Behavioral Intelligence. *International Journal of Neuroscience*, *17*(3), 183–191. <https://doi.org/10.3109/00207458208985922>
- 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>
- Schlüter, H., & Bermeitinger, C. (2017). Emotional Oddball: A Review on Variants, Results, and Mechanisms. *Review of General Psychology*, *21*(3), 179–222. <https://doi.org/10.1037/gpr0000120>

- Schneider, D. M., Sundararajan, J., & Mooney, R. (2018). A cortical filter that learns to suppress the acoustic consequences of movement. *Nature*, *561*(7723), 391–395.
<https://doi.org/10.1038/s41586-018-0520-5>
- Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences*, *11*(8), 349–355.
<https://doi.org/10.1016/j.tics.2007.06.005>
- 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>
- Seth, A. K. (2014). A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive Neuroscience*, *5*(2), 97–118. <https://doi.org/10.1080/17588928.2013.877880>
- Seth, A. K., & Hohwy, J. (2021). Predictive processing as an empirical theory for consciousness science. *Cognitive Neuroscience*, *12*(2), 89–90.
<https://doi.org/10.1080/17588928.2020.1838467>
- Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin*, *136*(6), 943–974. <https://doi.org/10.1037/a0020541>
- Šimundić, A.-M. (2013). Bias in research. *Biochemia Medica*, *23*(1), 12–15.
<https://doi.org/10.11613/BM.2013.003>
- Sommer, M. A., & Wurtz, R. H. (2008). Brain Circuits for the Internal Monitoring of Movements. *Annual Review of Neuroscience*, *31*(1), 317–338.
<https://doi.org/10.1146/annurev.neuro.31.060407.125627>
- 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*(1), 149–157. <https://doi.org/10.1007/s00221-012-3204-7>

- Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 38(4), 387–401. [https://doi.org/10.1016/0013-4694\(75\)90263-1](https://doi.org/10.1016/0013-4694(75)90263-1)
- 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>
- Thompson, S. C., Nierman, A., Schlehofer, M. M., Carter, E., Bovin, M. J., Wurzman, L., Tauber, P., Trifskin, S., Marks, P., Sumner, J., Jackson, A., & Vonasch, A. (2007). How Do We Judge Personal Control? Unconfounding Contingency and Reinforcement in Control Judgments. *Basic and Applied Social Psychology*, 29(1), 75–84. <https://doi.org/10.1080/01973530701331189>
- 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>
- Todorovic, A., Ede, F. van, Maris, E., & Lange, F. P. de. (2011). Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. *Journal of Neuroscience*, 31(25), 9118–9123. <https://doi.org/10.1523/JNEUROSCI.1425-11.2011>
- Todorovic, A., & Lange, F. P. de. (2012). Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields. *Journal of Neuroscience*, 32(39), 13389–13395. <https://doi.org/10.1523/JNEUROSCI.2227-12.2012>
- van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, 111(40), 14332–14341. <https://doi.org/10.1073/pnas.1402773111>

- Von Helmholtz, H. (n.d.). (1878/1971). The facts of perception. In *The Selected Writings of Hermann von Helmholtz*. R.Kahl, Ed. 115–185. Middletown, Wesleyan University Press.
- Walsh, K. S., McGovern, D. P., Clark, A., & O’Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. <https://doi.org/10.1111/nyas.14321>
- Wang, J., Mathalon, D. H., Roach, B. J., Reilly, J., Keedy, S. K., Sweeney, J. A., & Ford, J. M. (2014). Action planning and predictive coding when speaking. *NeuroImage*, *91*, 91–98. <https://doi.org/10.1016/j.neuroimage.2014.01.003>
- Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011a). The Self in Social Interactions: Sensory Attenuation of Auditory Action Effects Is Stronger in Interactions with Others. *6*(7), e22723. <https://doi.org/10.1371/journal.pone.0022723>
- Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011b). 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>
- Woodward, N. D., Karbasforoushan, H., & Heckers, S. (2012). Thalamocortical Dysconnectivity in Schizophrenia. *American Journal of Psychiatry*, *169*(10), 1092–1099. <https://doi.org/10.1176/appi.ajp.2012.12010056>
- Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122. <https://doi.org/10.1016/j.brainres.2009.06.032>

