WILEY

# Sensory suppression and increased neuromodulation during actions disrupt memory encoding of unpredictable self-initiated stimuli

Nadia Paraskevoudi<sup>1,2</sup> | Iria SanMiguel<sup>1,2,3</sup>

<sup>1</sup>Institut de Neurociències, Universitat de Barcelona, Barcelona, Spain

<sup>2</sup>Brainlab-Cognitive Neuroscience Research Group, Departament de Psicologia Clinica i Psicobiologia, University of Barcelona, Barcelona, Spain

<sup>3</sup>Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain

#### Correspondence

Iria SanMiguel, Brainlab-Cognitive Neuroscience Research Group, Departament de Psicologia Clinica i Psicobiologia, University of Barcelona, P. Vall d'Hebron 171, Barcelona 08035, Spain.

Email: isanmiguel@ub.edu

#### **Funding information**

MCIN/AEI/10.13039/501100011033, Grant/Award Number: MDM-2017-0729-18-2M Maria de Maeztu Center of Exce; MCIN/ AEI/10.13039/501100011033 and by "ERDF A way of making Europe", Grant/Award Number: PSI2017-85600-P; MCIN/ AEI/10.13039/501100011033 and by "ESF Investing in your future", Grant/ Award Number: RYC-2013-12577; Secretaria d'Universitats i Recerca de la Generalitat de Catalunya and the European Social Fund, Grant/ Award Number: FI-DGR 2019; Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya, Grant/ Award Number: Excellence Research Group 2017SGR-974

#### Abstract

Actions modulate sensory processing by attenuating responses to self- compared to externally generated inputs, which is traditionally attributed to stimulusspecific motor predictions. Yet, suppression has been also found for stimuli merely coinciding with actions, pointing to unspecific processes that may be driven by neuromodulatory systems. Meanwhile, the differential processing for self-generated stimuli raises the possibility of producing effects also on memory for these stimuli; however, evidence remains mixed as to the direction of the effects. Here, we assessed the effects of actions on sensory processing and memory encoding of concomitant, but unpredictable sounds, using a combination of selfgeneration and memory recognition task concurrently with EEG and pupil recordings. At encoding, subjects performed button presses that half of the time generated a sound (motor-auditory; MA) and listened to passively presented sounds (auditory-only; A). At retrieval, two sounds were presented and participants had to respond which one was present before. We measured memory bias and memory performance by having sequences where either both or only one of the test sounds were presented at encoding, respectively. Results showed worse memory performance - but no differences in memory bias -, attenuated responses, and larger pupil diameter for MA compared to A sounds. Critically, the larger the sensory attenuation and pupil diameter, the worse the memory performance for MA sounds. Nevertheless, sensory attenuation did not correlate with pupil dilation. Collectively, our findings suggest that sensory attenuation and neuromodulatory processes coexist during actions, and both relate to disrupted memory for concurrent, albeit unpredictable sounds.

#### K E Y W O R D S

auditory processing, EEG, memory, N1, pupillometry, self-generation

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. © 2022 The Authors. *Psychophysiology* published by Wiley Periodicals LLC on behalf of Society for Psychophysiological Research.

## **1** | INTRODUCTION

Forming predictions about upcoming events in the environment is crucial for all behaving organisms. A critical instance of such predictive processing is our ability to anticipate the sensory consequences of our own actions, which is essential for building a sense of self and shapes our perception of sense of agency (Gallagher, 2000). Although predictions have been suggested to facilitate perceptual processing in the wider sensory literature, in the action literature most studies report that the processing of predicted self-produced stimuli is attenuated (Press et al., 2020; Schröger et al., 2015), with only a few exceptions showing the opposite effect (e.g., Eliades & Wang, 2003; Reznik et al., 2014). Thus, several lines of research have shown that actions suppress the processing of the self-generated reafferent input (so-called selfgeneration effect) in a wide range of species (Chagnaud et al., 2015; Kelley & Bass, 2010; Kim et al., 2015; Requarth & Sawtell, 2011; Roy & Cullen, 2001; Schneider et al., 2014) and irrespective of sensory modality (auditory; Baess et al., 2011; Horváth, 2013a; Horváth, 2013b; Klaffehn et al., 2019; Martikainen et al., 2004; Mifsud & Whitford, 2017; SanMiguel et al., 2013; Saupe et al., 2013; Schafer & Marcus, 1973; Timm et al., 2013, visual; Hughes & Waszak, 2011; Mifsud et al., 2018; Roussel et al., 2013, 2014, and tactile; Blakemore et al., 1998; Hesse et al., 2010; Kilteni et al., 2020). Nevertheless, the exact mechanisms underlying the suppression of sensory responses to selfinitiated stimuli is still a matter of debate (for reviews see Horváth, 2015; Hughes et al., 2013; Schröger et al., 2015). Interestingly, beyond modulating sensory responses, selfgeneration also appears to have consequences for memory encoding. The so-called "production effect" (Brown & Palmer, 2012; MacDonald & MacLeod, 1998) refers to memory benefits reported for stimuli that are self-generated in a predictive context; f example, it is easier to remember a piano melody that was learnt by playing versus listening to it (Brown & Palmer, 2012). Nevertheless, the bulk of the evidence for the production effect on memory comes from behavioral studies, and thus the underlying neurophysiological mechanisms remain largely unexplored. Crucially, given that memory relies on the sensory representation (e.g., Nyberg et al., 2000; Wheeler et al., 2000), the production effect could be a direct consequence of the differential sensory processing of self-initiated stimulation. However, to date, the possible relationship between the effects of self-initiation on sensory processing and memory has not been investigated. Here we sought out to bridge the gap between these two different research lines that have evolved separately over the last decades, aiming to identify the possibly shared neurophysiological mechanisms involved in each of these effects, focusing on the auditory modality.

In the following paragraphs, we summarize findings that have inspired the present work, in an attempt to highlight the need of examining in detail the processes driving the self-generation effects and their possible links with the encoding of self-generated stimulation in memory.

# **1.1** | Sensory processing of selfinitiated stimuli

To date, most studies assessing the effects of actions on sensory processing, have attributed the attenuation effects to a predictive mechanism that predicts the sensory consequences of our actions (e.g., Bays et al., 2006; Blakemore et al., 1998; Martikainen et al., 2004). This view was inspired to a great extent by classic physiology research and the reafference principle (Sperry, 1950; von Holst & Mittelstaedt, 1950), and later by motor control theories that have further refined this idea by suggesting that sensory attenuation is an integral part of our motor abilities (Miall & Wolpert, 1996; Wolpert et al., 1995). This line of work was the first to point to the involvement of forward and inverse models in sensorimotor behavior: The former estimate the future state of the system by comparing the predicted to the actual sensory consequences of the action, while the latter allow the system to estimate a motor plan (and its associated motor commands) so as to achieve a desired state (Miall & Wolpert, 1996; Wolpert et al., 1995). Especially forward models have been at the core of the dominant cancellation account that has been widely used to explain the self-generation effects (also known as the comparator model; Blakemore et al., 1998; Frith et al., 2000; Wolpert & Flanagan, 2001). According to this account, the suppression effects result from the operation of a forward model that generates stimulus-specific prediction signals before or during an action and sends them from the motor to the corresponding sensory cortices (Sperry, 1950; von Holst & Mittelstaedt, 1950). These motor-induced predictions of sensory reafference (i.e., corollary discharge) are compared to the sensory input generated by one's actions, and only the difference between the two (i.e., prediction error) is sent to higher stages of the neuronal hierarchy for further processing, ultimately suppressing the processing of the anticipated event in order to prioritize the most informative unexpected inputs (Friston, 2005). An important implication of the cancellation model is that the effects should be specific to the predicted stimulus, and thus mediated by sensory-specific areas (i.e., the effect should reflect attenuation of the predicted stimulus' representation in the sensory-specific areas).

In fact, there is mounting evidence showing that the attenuation effects for self-generated stimuli are (at least partly) stimulus-specific (Aliu et al., 2009; Fu et al., 2006;

Hashimoto & Sakai, 2003; Heinks-Maldonado et al., 2005; Houde et al., 2002; Martikainen et al., 2004; Ott & Jäncke, 2013). Most studies supporting the specificity of the effects have employed the contingent self-generation paradigm, where neural responses to sounds generated by the participants in a fully predictable fashion are compared to the responses elicited by externally generated sounds (e.g., Baess et al., 2009, 2011; Martikainen et al., 2004; Mifsud & Whitford, 2017) and have shown attenuated auditory N1 and P2 event-related potential (ERP) amplitudes (for a review see Schröger et al., 2015). Crucially, suppression is larger when the match between predicted and actual sensory input is more precise (Baess et al., 2008; Fu et al., 2006; Hashimoto & Sakai, 2003; Heinks-Maldonado et al., 2005; Houde et al., 2002), and it is suggested to occur within the auditory cortex (Aliu et al., 2009; Flinker et al., 2010; Martikainen et al., 2004), providing further support to the stimulus-specificity of the effects.

However, additional, stimulus-unspecific processes are also known to modulate sensory and perceptual responses during actions (Korka et al., 2021; Press et al., 2020; Press & Cook, 2015; Schröger et al., 2015). For example, there is evidence for generalized unspecific attenuation during movements (e.g., saccadic suppression and somatosensory gating, Chapman et al., 1987; Cohen & Starr, 1987; Crapse & Sommer, 2008; Ross et al., 2001; Williams & Chapman, 2000, 2002; Williams et al., 1998), which suggests that during actions the system might expect some action-related consequence, without necessarily generating a specific sensory prediction regarding the effect of the motor act. Supporting evidence to this idea comes also from studies showing suppression of responses when the stimulus merely coincides with an action (Hazemann et al., 1975; Horváth, 2013a, 2013b; Horváth et al., 2012; Makeig et al., 1996; Tapia et al., 1987), that is in the absence of a predictive relationship between action and stimulus. A recent study further examined the specificity of the attenuation effects, by assessing whether the self-generation effects (measured as auditory N1 attenuation) reflect a genuine modulation within the auditory cortex (SanMiguel et al., 2013). Based on evidence showing that the N1 response reflects the overlap of several components (Näätänen & Picton, 1987), SanMiguel et al. (2013) assumed that if the attenuation effects reflect stimulus-specific prediction mechanisms, then the effects should be observable in the sensory-specific components, namely the N1 at the mastoids (i.e., generated by tangentially oriented sources in the auditory cortex) and the "T complex" (i.e., the first and second negative peaks, known as Na and Tb, identified on anterior temporal sites, and generated by radial sources in the superior temporal gyrus; Tonnquist-Uhlen et al., 2003; Wolpaw & Penry, 1975). However, they showed that sensory attenuation mainly reflects the modulation of the unspecific N1 component,

PSYCHOPHYSIOLOGY SPR

which is suggested to be the cortical projection of a reticular process facilitating motor activity, related to the orienting response (Näätänen & Picton, 1987). In contrast, they could not find a clear attenuation of the specific N1 components (cf. Horváth et al., 2012; Timm et al., 2013). Collectively, the findings reviewed this far point to a complex picture of possibly coexisting specific and unspecific effects of actions on sensory responses and suggest that the effects cannot be fully attributed to stimulus-specific prediction mechanisms.

As we have seen so far, converging evidence suggests that stimulus-unspecific processes might partly drive the sensory attenuation effects, thereby raising the need to identify the mechanism driving the unspecific attenuation during movement. One intriguing possibility is that the suppression effects may be mediated by a halo of neuromodulation surrounding actions, which would unspecifically gate auditory processing for stimuli presented close in time with the motor act. Neuromodulatory influences on the action-induced suppression effects seem plausible considering that rodent studies show that actions trigger a cascade of neuromodulatory processes (Eggermann et al., 2014; McGinley et al., 2015; Vinck et al., 2015), and that motor and neuromodulatory inputs overlap in auditory areas during movement (Nelson & Mooney, 2016; for a review see Schneider & Mooney, 2018). A possible candidate for creating a halo of neuromodulation that could mediate unspecific effects during movement could be the locus coeruleus norepinephrine system (LC-NE). This possibility has received substantial support recently, mainly by data showing a close association between pupil diameter - a proxy of LC-NE activity - (Aston-Jones & Cohen, 2005; Joshi et al., 2016; Murphy et al., 2014; Vinck et al., 2015) and actions in rodents (Lee & Margolis, 2016; McGinley et al., 2015; Vinck et al., 2015), monkeys (Bornert & Bouret, 2021), and humans (Lubinus et al., 2021; Yebra et al., 2019). However, to the best of our knowledge, there have been no attempts to test for possible links between sensory attenuation for self-generated sounds and neuromodulation (i.e., as reflected in pupil diameter) during actions.

# **1.2** | Memory encoding of selfinitiated stimuli

Meanwhile, actions might also affect other high-level processes beyond the immediate sensory processing. Strikingly, despite the mounting evidence pointing to a differential processing of self- and externally generated stimuli (e.g., for a review see Schröger et al., 2015), but also to modulatory effects of movements on areas supporting memory processes (e.g., hippocampal and parahippocampal activity; Halgren, 1991; Mukamel et al., 2010; Rummell et al., 2016), there have been only

few attempts to assess the effects of actions on memory encoding of self-generated stimulation. One research line - known as the "production effect" - has shown improved memory for self-produced stimuli compared to their passively listened comparisons (e.g., spoken words or played melodies compared to passively listened ones; Brown & Palmer, 2012; MacDonald & MacLeod, 1998), which has been attributed to the increased distinctiveness afforded by the extra mnemonic information of having produced these items that is not present for silently read words (Conway & Gathercole, 1987; Mama & Icht, 2016; Ozubko et al., 2012). This line of work, however, contrasts with the predictions made by predictive coding theories of memory. According to this account, memory is driven by the amount of surprise (i.e., prediction error) associated with an item, such that items eliciting larger prediction error responses (as reflected in larger evoked potentials or fMRI signal) should be encoded better in memory than the less surprising, predictable ones (Bar, 2009; Greve et al., 2017; Heilbron & Chait, 2018; Henson & Gagnepain, 2010; Krawczyk et al., 2017; Pine et al., 2018). This framework would, therefore, predict memory enhancements for the externally generated sounds in a typical contingent paradigm where they inherently elicit larger prediction error (and enhanced sensory responses) compared to the more predictable self-generated stimuli.

## **1.3** | The present study

To the best of our knowledge, there have been no attempts to simultaneously address the effects of selfgeneration on sensory processing and memory encoding of sounds and assess the possible relationship between these two phenomena and their underlying neurophysiological mechanisms. Based on the evidence indicating that self-generation effects might not be solely attributed to stimulus-specific motor predictions (e.g., Horváth et al., 2012; SanMiguel et al., 2013), we hypothesize that performing an action may trigger the activation of stimulus-unspecific neuromodulatory mechanisms, namely the LC-NE system. We hypothesize that this motor-driven noradrenergic activity may modulate the processing of sounds presented during the performance of the action, leading to suppressed sensory responses and altered memory encoding as a consequence of the latter.

In order to test these hypotheses, here we examine whether and how motor actions affect sensory processing and memory encoding of concomitant, but unpredictable sounds, by employing a combination of a self-generation and memory recognition task, while monitoring the brain's and the pupil's responses to sounds that are either presented passively or that coincide in time with a motor act. The aim of this study is twofold: Our first aim is to investigate the role of the neuromodulatory LC-NE system in the motor-driven modulation of auditory processing of self-generated sounds. Related to this first aim, we have specific hypotheses about the effects of actions on sensory responses and pupil diameter. As for the sensory responses, we hypothesize that event-related potentials (i.e., N1 at vertex but not at the mastoids, P2, and Tb) should be attenuated for sounds coinciding with an action (cf. Hazemann et al., 1975; Horváth, 2013a, 2013b; Horváth et al., 2012; Makeig et al., 1996). As for the pupil diameter, we hypothesize that neuromodulatory activity (i.e., reflected in pupil diameter; Aston-Jones & Cohen, 2005) should increase during actions (cf. Lee & Margolis, 2016; McGinley et al., 2015; Simpson, 1969; Vinck et al., 2015; Yebra et al., 2019) and that it should correlate with the sensory attenuation effects measured on the auditory eventrelated potentials. Our second aim is to assess whether the differential sensory processing of stimuli paired with an action affects their encoding in memory. We expect to observe differences in memory performance between passively encoded sounds and sounds that coincide with an action at encoding. However, given the lack of contingency between actions and sounds in the present paradigm as compared to the typical production effect studies, as well as the mixed evidence (memory for self-initiated stimulation is enhanced in previous production effect studies but should be reduced according to predictive coding views; Brown & Palmer, 2012; Henson & Gagnepain, 2010; MacDonald & MacLeod, 1998), we remain agnostic as to the direction of the effect. However, critically, we hypothesize that the potential differences in the memory encoding of sounds presented with or without a concomitant action should be driven by, and thus correlate with, the differential neurophysiological responses (i.e., event-related potentials and pupil diameter) at encoding for sounds that were either paired with an action or not.

# 2 | METHOD

All the scripts for the experimental stimulation and data analysis are available on Open Science Framework, along with the detailed experimental protocol (https://osf.io/238xe/?view\_only=4b6d8fdc2a2f4982bac76a72d c78e0ec).

## 2.1 | Participants

Twenty-six healthy, normal-hearing subjects, participated in the present study. Participants were typically undergraduate university students at the University of Barcelona. Data from three participants had to be excluded due to technical problems, inability to comply with the task instructions, or excessive artifacts in the EEG recording, leaving data from 23 participants (6 men, 17 women,  $M_{age} = 24.82$ , age range: 21–36). None of them had any hearing impairments, suffered, or had suffered from psychiatric disorders or had taken substances affecting the central nervous system 48 hr prior to the experiment. All participants gave written informed consent for their participation after the nature of the study was explained to them and they were monetarily compensated (10 euros per hour). Additional materials included a personal data questionnaire, a data protection document, and five personality trait questionnaires. The study was accepted by the Bioethics Committee of the University of Barcelona.

# 2.2 | General experimental design

Each trial consisted of three phases: the encoding phase, the retention phase, and the retrieval phase (Figure 1).

# 2.2.1 | Encoding phase

At the start of each trial, subjects were presented with a row of six vertical lines on the screen, separated in semirandom distances from each other. The positions of vertical lines were distributed based on the sequence presented in each trial. During the whole duration of the encoding period (12s), a horizontal line moved at a stable pace across the screen from left to right, intersecting each of the vertical lines as it advanced. Participants pressed a button with their right thumb every time the horizontal line reached one of the vertical ones. Only half of these presses produced a sound (Motor-auditory event; MA). The other half did not result in the presentation of a sound (Motoronly event; M). Additionally, three more sounds were presented passively to the participants without being triggered by a button press (Auditory-only event; A). Thus, in every trial, the encoding set consisted of six different sounds to be remembered, delivered within nine events (three motor-only (M), three Motor-auditory sounds (MA), and three Auditory-only (A) sounds). The interval

PSYCHOPHYSIOLOGY



**FIGURE 1** Schematic representation of the experimental design, showing an example trial for the two types of sequences employed: 2 T sequences (left) and 1 T sequences (right). Each trial consisted of three phases: Encoding, retention, and retrieval. The two types of sequences differed only in the retrieval phase. The different boxes represent the visual stimulation as a function of time. Each trial started with the encoding phase, where six vertical lines were initially presented (top box), and subsequently a horizontal line started moving across the screen from left to right, intersecting each of the six vertical lines as it advanced. Participants were asked to press a button every time the horizontal line reached one of the vertical ones. Only half of these presses produced a sound (motor-auditory; MA). The other half did not result in the presentation of a sound (motor-only; M). Additionally, three more sounds were presented passively to the participants without being triggered by a button press (auditory-only condition; A). Therefore, six different sounds (shown by the different colors of the sounds in the figure) were presented during encoding and had to be maintained in memory for a 3 s retention period (box with fixation cross). In the retrieval phase, arterieval were both presented at encoding, one encoded as MA and the other encoded as A. in the 1 T sequences (right), only one of the two sounds was presented at encoding that was either encoded as A or encoded as MA (in the figure an "Encoded as MA" sound is shown), while the other sound was not presented at encoding (new). After the presentation of the test sounds, a question mark appeared on the screen, prompting participants to respond whether the first or the second test sound was presented during the encoding phase.

5 of 25

between any two events (MA, M or A) varied from 0.8 to 2.4 s, in steps of 0.05 s, while the interval between any two sounds varied between 1.6 and 2.4 s in steps of 0.05. The encoding phase finished when the horizontal line had intersected all the vertical ones and arrived at the right of the screen. If the task was performed correctly (i.e., all required button presses were performed), the trial continued into the retention phase. Otherwise, an error message appeared on the screen indicating that the participant did not press the button every time the horizontal line reached a vertical one, and a new trial began.

#### 2.2.2 | Retention phase

During the subsequent retention phase, participants were presented with a fixation cross on the screen for 3 s and they were asked to remember the six individual sounds that had been presented in the encoding phase.

## 2.2.3 | Retrieval phase

In the retrieval phase, participants were presented with two test sounds with a 2 s sound-to-sound onset asynchrony (indicated by the visual stimuli "Sound 1" and "Sound 2," respectively). Subsequently, a question mark appeared on the screen, prompting participants to respond whether the first or the second test sound was presented during the encoding phase. We employed two different types of sequences (see Section 2.3.1. Sequences for more details) that differed only in this retrieval phase: "Two Test Sounds at Encoding" sequences (henceforth 2 T; Figure 1, left panel) and "One Test Sound at Encoding" sequences (henceforth 1 T; Figure 1, right panel). Unbeknownst to the participants, in the 2 T sequences, both test sounds had been presented at encoding, while in the 1 T sequences only one of the test sounds had been presented at encoding. Nevertheless, participants made a forced choice between the two sounds on every trial. The response window was 2 s. After the end of the response window, a fixation was presented for 2 s (inter-trial interval) before the start of the next trial.

#### 2.3 | Stimuli

#### 2.3.1 | Sequences

Two types of sequences were created, differing in whether both or only one of the test sounds presented at retrieval were also present during the encoding phase. In the "Two Test Sounds at Encoding" (2 T; Figure 1, left panel) and unbeknownst to the participants, the two test sounds presented passively at retrieval were also presented in the encoding sequence, one as a motorauditory (Encoded as MA) and the other one as an auditory-only event (Encoded as A). These sequences were intended to measure memory bias. In the "One Test Sound at Encoding" (1 T; Figure 1, right panel), only one of the test sounds at retrieval was presented at encoding, either as an MA (Encoded as MA) or as an A event (Encoded as A), while the other sound was not presented at encoding (New sound). The 1 T sequences were intended to measure memory performance. They were introduced only for the behavioral data and were not used for the EEG and pupillometry analyses. This design allowed us to have enough trials for Encoded as A and Encoded as MA sounds at retrieval, keep the experiment's duration within a reasonable time, and obtain an additional objective measure of memory performance in the 1 T sequences besides the measure of memory bias obtained in the 2 T sequences. Five of the 1 T sequences were randomly chosen to be used during the practice block.

Importantly, the same sounds in the same encoding sequence positions were used as either A or MA in different trials, which allowed us to compare between physically identical auditory sequences that only differed in the actions performed. Additionally, we counterbalanced the order of the sounds at encoding that would be later used as test at retrieval for the 2 T sequences, the order of the two retrieval sounds, and the position of the test sounds in the encoding sequence. Related to the latter, we discarded the first and last position of the encoding sequence for placing test sounds to avoid primacy and recency effects, which refer to an improvement in memory retention for stimuli that have been presented first or last in a list, respectively (Mondor & Morin, 2004). However, we included 20 catch trials with test sounds in those positions, which were randomly interleaved with the experimental sequences described above and discarded from all analyses.

## 2.3.2 | Auditory stimuli

For the main experiment, 255 different, environmental, natural, complex, and non-identifiable sounds were gathered from the libraries of McDermott Sound Library (http://mcdermottlab.mit.edu/svnh/Natural-Sound/ Stimuli.html) and Adobe (https://offers.adobe.com/en/ na/audition/offers/audition\_dlc.html). These sounds were then edited to all have 250 ms of duration, a sampling rate of 44.1 kHz and to be played at 16 bits, mono and with 70 dB of intensity. Subsequently, six volunteers that did not participate in the main experiment rated the 255 sounds based on their identifiability (i.e., how easy it was to assign a name to them). All sounds were presented to them in a randomized order and each sound was presented twice. The volunteers rated them in a scale from 1-3 (1 = identifiable, 2 = not sure, 3 = not identifiable), and the mean score for each sound was calculated. The 108 less identifiable sounds were selected to construct the unique experimental sound sequences. The sounds used in the practice block consisted of 35 pure tones of different frequencies, ranging from 300 Hz to 3700 Hz in steps of 100.

# 2.4 | Apparatus

The visual stimuli were presented on an ATI Radeon HD 2400 monitor. The auditory stimuli were presented via the Sennheiser KD 380 PRO noise canceling headphones. To record participants' button presses at encoding (right hand button press) and behavioral responses at retrieval (left hand button presses), we used the Korg nanoPAD2. The buttons of this device do not produce any mechanical noise when pressed, and, thus, do not interfere with our auditory stimuli. The presentation of the stimuli and recording of participants' button presses and responses were controlled using MATLAB R2017a, the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007), and the Eyelink add-in toolbox for eyetracker control.

EEG activity was acquired with Neuroscan 4.4 software and Neuroscan SynAmps RT amplifier (NeuroScan, Compumedics, Charlotte, NC, USA). We recorded continuously with Ag/AgCl electrodes from 64 standard locations according to the 10% extension of the International 10-20 system (Chatrian et al., 1985; Oostenveld & Praamstra, 2001) mounted in a nylon cap (Quick-Cap; Compumedics, Charlotte, NC, USA). An additional electrode was placed at the tip of the nose (serving as online reference). The vertical electrooculogram (EOG) was measured with two electrodes placed above and below the left eye, and the horizontal EOG with two electrodes placed on the outer canthi of the eyes referenced to the common reference (unipolar montage). The ground electrode was placed at AFz. All impedances were kept below 10 k $\Omega$ during the whole recording session and data were sampled at 500 Hz.

Concurrently with the EEG recording, horizontal and vertical gaze position, as well as the area of the pupil, were recorded using EyeLink 1000 desktop mount (SR Research, sampling rate: 1000Hz; left eye recordings except for three participants for whom the right eye was recorded instead). The pupil was assessed in the centroid mode of the eye tracker, which uses a center-of-mass

algorithm. This algorithm detects the pupil area by identifying the number of black pixels and its center on the video image. Importantly, in contrast to methods using ellipse fitting for the measurement of the pupil, this method is hardly affected by noise (S-R Research Eyelink-CL Manual, p. 71).

## 2.5 | Procedure

Prior to the start of the experiment, participants were asked to complete several questionnaires and were given written and verbal instructions about the task. Specifically, they were told that at the start of every trial they would first see six vertical lines and that a horizontal line would start to move from left to right, intersecting each vertical line as it advanced. They were explicitly instructed to press the predefined button every time the horizontal line crossed one of the vertical ones (not too early / late and no more than one button press per vertical line). They were told that they would hear several sounds while the line advanced, some of them might coincide with the button press and some of them not and that they should try to memorize all the sounds presented because later they would be tested in memory. Finally, they were told that once the horizontal line had crossed all the vertical lines, a fixation cross would appear and subsequently two sounds would be presented (indicated by the visual stimuli "Sound 1" and "Sound 2," respectively) and that they would have only 2 s to reply which one of the two sounds was presented during the encoding period of the trial. They were asked to make a choice on each trial between the two test sounds within the response window.

After explaining them the instructions, participants were seated in an electrically and acoustically shielded room and were asked to place their head on a chinrest at approximately 60 cm from the screen. Eyetracker calibration was performed first at the start of the experiment and then every six blocks. In order to familiarize themselves with the task, participants first completed a practice block of 5 trials and repeated it as many times as they needed to make sure they understood how to perform the task. During the main experiment, participants completed a total of 236 trials: 56 1 T trials, 160 2 T trials and 20 catch trials. These were divided in 24 blocks. 20 of them comprised of 10 trials (9 experimental and 1 catch trial) and the remaining 4 comprised of 9 trials (all of them experimental trials). At the end of each block, a message appeared informing participants about the number of errors (i.e., not pressing the button when required) and extra-presses (i.e., more than the required button presses) at the encoding phase, as well as the number of missed responses at retrieval for this block.

Catch trials, as well as trials including errors in buttonpressing and missed responses were discarded from further analyses. Participants took a break of approximately 5 minutes every six blocks to prevent fatigue. The experiment lasted for approximately 1.5 hours excluding the EEG preparation.

#### 2.6 Data analysis

### 2.6.1 | Behavioral analysis

To test for differences in memory bias (2 T sequences) and memory performance (1 T sequences) for sounds encoded as A or MA, we performed two different analyses. For the 1 T sequences, we calculated the percent correct for the sounds at retrieval (i.e., memory performance), separately for those that were Encoded as MA and Encoded as A, which was subsequently submitted to a two-sided paired samples t test. For the 2 T-trial sequences, we calculated the percent recall for sounds Encoded as MA and Encoded as A and tested for differences in memory bias, using a two-sided paired samples t test. We complemented the frequentist t tests with corresponding Bayesian t tests, separately for the 1 and 2 T sequences. For both Bayesian comparisons, the Bayes factor  $(BF_{10})$  for the alternative hypothesis (i.e., the difference of the means is not equal to zero) was calculated (using the function ttestBF of the BayesFactor package in R). Specifically, the null hypothesis was specified as a point-null prior, corresponding to a standardized effect size  $\delta = 0$ , and the alternative hypothesis was defined as a Cauchy prior distribution centered around 0 with a scaling factor of r = .707 (Rouder et al., 2009). In line with the Bayes Factor interpretation (Lee & Wagenmakers, 2013) and with previous work reporting Bayes Factors (Korka et al., 2019, 2020; Marzecová et al., 2018), data were taken as moderate evidence for the alternative hypothesis if the  $BF_{10}$  was greater than 3, while values close to 1 were considered only weakly informative. Values greater than 10 were considered strong evidence for the alternative (or null) hypothesis.

#### 2.6.2 | EEG preprocessing

EEG data were analyzed with EEGLAB (Delorme & Makeig, 2004) and plotted with EEProbe (ANT Neuro). Data were high-pass filtered (0.5 Hz high-pass, Kaiser window, Kaiser  $\beta$  5.653, filter order 1812), manually inspected so as to reject atypical artifacts and identify malfunctioning electrodes, and corrected for eye movements with Independent Component Analysis, using

the compiled version of runica (binica) that uses the logistic infomax ICA algorithm (Onton & Makeig, 2006). Components capturing eye movement artifacts were rejected by visual inspection and the remaining components were then projected back into electrode space. Data were then low-pass filtered (30 Hz low-pass, Kaiser window, Kaiser  $\beta$  5.653, filter order 1812), remaining artifacts were rejected by applying a 75 µV maximal signal-change per epoch threshold, and malfunctioning electrodes were interpolated (spherical interpolation). A -100 to +500 ms epoch was defined around each event both at the encoding and the retrieval phase. The data were subsequently baseline corrected (100 ms prior to the event). We calculated the average wave for each event of interest, as well as the grand average for the whole sample. Specifically, we obtained the averages for the MA, A, and M events at encoding, while for the retrieval data, we binned the responses to motor-auditory and auditory-only sounds as a function of memory (i.e., Encoded as MA and Encoded as A at retrieval that were remembered or forgotten). For each condition of interest the number of remaining trials used for the analyses after trial rejection were: Auditory-only (M = 424.9, SD = 46.9), Motor-auditory (M = 427.2, SD = 40.6), Motor-only (M = 429, SD = 40.8), Encoded as A and forgotten (M = 68, SD = 11.7), Encoded as A and remembered (M = 64, SD = 14.7), Encoded as MA and forgotten (M = 64.1, SD = 14.2), Encoded as MA and remembered (M = 67.7, SD = 11.9).

To assess self-generation effects at encoding, MA sound responses were corrected for motor activity by subtracting the motor-only (M) averages from the motor-auditory (MA) averages, as the signal obtained in the MA condition represents the brain signal elicited by the sound, but also by the planning and execution of the finger movement to press the button. We, therefore, obtained a motor-corrected wave that only included the brain signal elicited by the MA sound. Self-generation effects at encoding were then assessed by comparing responses to MA sounds corrected for motor activity (MA-M) with the responses elicited by the auditory-only sounds (A). Self-generation effects are presented in all figures as the difference wave between the motor-auditory (corrected) sound responses and the auditory-only sound responses (A-[MA-M]). No motor correction was performed at retrieval as both test sounds were presented passively.

### 2.6.3 | ERP analysis

For all the effects of interest at encoding, we examined responses separately for the N1 and P2 at Cz (N1, P2)

and at the mastoids (henceforth,  $N1_{mast}$ ,  $P2_{mast}$ ), the P3 component at Pz, and the N1 subcomponents Na and Tb at T7 and T8. The same components except for P3 were examined at retrieval. The windows were defined after visual inspection of the data by locating the highest negative or positive (depending on the component of interest) peak in the usual latencies for each component as reported by previous work (SanMiguel et al., 2013). Specifically, time windows for N1 (and N1<sub>mast</sub>), P2 (and P2<sub>mast</sub>), Na, and Tb were defined on the grand-averaged waveforms of the auditory-only sounds as previously reported (e.g., SanMiguel et al., 2013). Na and Tb were identified as the first and second negative peaks, respectively, identifiable after sound onset on electrodes T7 and T8, as recommended by Tonnquist-Uhlen et al. (2003). N1/N1<sub>mast</sub> and P2/P2<sub>mast</sub> were identified as the negative and positive peaks occurring in the window ~70 to 150 ms, and ~150 to 250 ms after stimulus onset on Cz, respectively, showing reversed polarity at the mastoid electrodes. P3 at encoding was identified as the peak of the difference wave (A -[MA-M]) in the P3 window range based on previous work (e.g., Baess et al., 2008). The time windows for the N1/N1<sub>mast</sub>, P2/P2<sub>mast</sub>, P3, Na, and Tb peaks were centered on the identified peaks  $\pm 13$ , 25, 15, 10, and 15 ms, respectively. Therefore, the final time windows used to calculate the average component amplitudes were the following: N1/N1<sub>mast</sub> 94-120 ms, P2/ P2<sub>mast</sub> 174-224 ms, P3 256-286 ms, Na 72-92 ms, Tb 120-150 ms. Given variations in peak latencies across the conditions, the width of the windows was defined such that it could capture the peak of the MA sound waveform as well, and it was proportional to the width of the component. For the encoding data, we performed paired samples t tests with the factor Sound Type (A vs. MA) to test for differences in N1, P2 and P3, and a repeated measures ANOVA with factors Sound Type (A vs. MA) x Laterality (M1 vs. M2 or T7 vs. T8) to test for differences in N1<sub>mast</sub>, P2<sub>mast</sub> and Na, Tb, respectively. For the retrieval data we performed  $2 \times 2$  ANOVAs with the factors Sound Type (Encoded as A vs. Encoded as MA) and Memory (Remembered vs. Forgotten) on N1 and P2, while for the N1<sub>mast</sub>, P2<sub>mast</sub>, Na, and Tb an additional factor Laterality was introduced in the ANOVAs (i.e., M1 vs. M2 or T7 vs. T8).

## 2.6.4 | Pupillometry analysis

Missing data and blinks, as detected by the EyeLink software, were padded by 100ms and linearly interpolated. Additional blinks were found using peak detection on the velocity of the pupil signal and linearly interpolated (Urai et al., 2017). Blinks separated by less than 250ms PSYCHOPHYSIOLOGY

9 of 25

were aggregated to a single blink. The interpolated pupil time series were bandpass filtered using a 0.05-4 Hz thirdorder Butterworth filter. Low-pass filtering reduces measurement noise not likely to originate from physiological sources, as the pupil functions as a low-pass filter on fast inputs (Binda et al., 2013; Hoeks & Levelt, 1993). Highpass filtering removes slow drifts from the signal that are not accounted for by the model in the subsequent deconvolution analysis. First, we estimated the effect of blinks and saccades on the pupil response through deconvolution and removed these responses from the data using linear regression using a procedure detailed in previous work (Knapen et al., 2016; Urai et al., 2017). The residual bandpass filtered pupil time series was used for the evoked analyses (Van Slooten et al., 2019). After zscoring per trial, we epoched trials (epoching window -0.5 to 1.5post-event), baseline corrected each trial by subtracting the mean pupil diameter 500 ms before onset of the event and resampled to 100 Hz.

For each participant, we first obtained the average evoked response for the main events of interest. Specifically, we obtained the averages for the A and MA events at encoding, while at retrieval we obtained the averages for the Encoded as A and Encoded as MA sounds, separately for the remembered and the forgotten ones. We used non-parametric permutation statistics to test for the group-level significance of the individual averages, separately for encoding and retrieval. Specifically, we computed *t* values of the difference between the two conditions of interest and thresholded these t values at a p value of .05. Each cluster was constituted by the samples that passed the threshold of the *p* value. The cluster statistics was chosen as the sum of the paired *t*-values of all the samples in the cluster. First, we compared the pupil response to motor-auditory and auditory-only events at encoding. For the retrieval data, we aimed to test for possible main effects of Sound Type (Encoded as A vs. Encoded as MA) and Memory (Remembered vs. Forgotten), as well as for possible interactions between the two. For the main effects of Sound Type and Memory at retrieval, the permutation statistics were performed between Encoded as A and Encoded as MA sounds (irrespective of their memory) and between Remembered and Forgotten sounds (irrespective of how they were encoded before), respectively. To test for possible interactions, the cluster-permutation test was performed on the difference waves ([Encoded as A and remembered - Encoded as MA and remembered] and [Encoded as A and forgotten – Encoded as MA and forgotten]). For each statistical test, this procedure was performed by randomly switching labels of individual observations between these paired sets of values. We repeated this procedure 10,000 times and computed the difference

between the group means on each permutation. The obtained p value was the fraction of permutations that exceeded the observed difference between the means (i.e., two-sided dependent samples tests). The pupil preprocessing and analysis was performed with custom software based on previous work (Urai et al., 2017) using Fieldtrip (Oostenveld et al., 2011).

### 2.6.5 | Correlations

Finally, we hypothesized that the electrophysiological and neuromodulatory effects at encoding (i.e., sensory suppression and pupil dilation for MA events) might be driving any memory encoding differences between A and MA sounds, and that neuromodulation might be behind the suppression of ERP responses to MA sounds. To assess these relationships, we tested for possible correlations between the behavioral, electrophysiological and neuromodulatory (i.e., pupil diameter) effects of actions. Only those differences between MA and A events that were found to be significant in the previous analyses were introduced in the correlation analyses. For all the behavioral and the electrophysiological effects, we first calculated the difference by subtracting the MA from A values (i.e., difference in memory and ERP amplitude for each component of interest between A and MA). Regarding the ERPs identified in two electrodes (e.g., Na, Tb, N1<sub>mast</sub>, P2<sub>mast</sub>), we calculated the mean amplitude across the two (T7/T8 and M1/M2, respectively). For the pupil data, we used the peak of the difference wave between A and MA events at encoding. We then submitted these values to a Pearson correlation coefficient to test for correlations between (a) the effects on ERPs at encoding and memory performance/bias (1 and 2 T sequences, respectively), (b) the neuromodulatory effects at encoding and memory performance/bias (1 and 2 T sequences, respectively), and (c) the effects on the ERPs and the neuromodulatory effects at encoding. In all correlations, for the ERPs, the larger the attenuation effects for the negative (N1, P2<sub>mast</sub>, Na, Tb) and positive (N1<sub>mast</sub>, P2, P3) components, the more negative and positive the values, respectively. Conversely, for the pupil and the behavioral data, the more negative the value, the larger the pupil diameter and the worse the memory performance for MA sounds.

## 3 | RESULTS

All statistical analyses were performed using R (version 3.6.0). For all the *t*-tests performed, we first confirmed that the assumption of normality was not violated (Shapiro–Wilk normality test p > .05). As we mentioned before

(see Method), the 1 T sequences were introduced to be used only for the behavioral analyses. For the EEG and pupil analyses, we only included the data from the 2 T sequences, after confirming that the results would remain the same when including the 1 T sequences as well.

## 3.1 | Behavioral performance

For the analysis of the behavioral data, we calculated the percent correct (i.e., memory performance in the 1 T sequences) and the percent recall (memory bias in the 2 T sequences) for sounds that were encoded as motor-auditory or auditory-only (see Figure 2). For the 1 T sequences, we obtained significantly better memory performance for sounds that were encoded as auditory-only compared to those that coincided with participants' motor acts in the previous encoding phase, t(22) = 3.15, p = .005, d = 0.66 $(M_{MA} = 0.757, M_A = 0.799, SD_{MA} = 0.108, SD_A = 0.0924).$ This difference, however, was not reflected in memory bias since we did not find significant differences between the Encoded as A and Encoded as MA sounds in the 2 T sequences, where both of the test sounds were presented at encoding, t(22) = 1.14, p = .267 (M<sub>MA</sub> = 0.509, M<sub>A</sub> = 0.491,  $SD_{MA} = 0.0395$ ,  $SD_A = 0.0395$ ). The absence of significant differences in memory bias may suggest that they remembered both sounds as evident by the generally high accuracy (i.e., mean performance in the 1 T sequences = 0.78with standard deviation of 0.1) which led them to choose randomly between A and MA sounds in 2 T sequences. We complemented the frequentist t tests with corresponding Bayesian t tests, separately for memory performance (1 T sequences) and memory bias (2 T sequences). The Bayesian t tests for the 1 and 2 T sequences yielded similar results as the ones obtained from the frequentist *t* tests. Specifically, this analysis brought strong evidence for the alternative hypothesis in the case of 1 T sequences ( $BF_{10} = 9.375$ ), while the Bayesian t test for the 2 T sequences, brought weak evidence for the alternative hypothesis ( $BF_{10} = 0.389$ ).

# 3.2 | Electrophysiological responses at encoding

Figure 3a shows all the studied peaks identified on the passive sound responses for the encoding conditions at the relevant electrodes for each peak. The motor-auditory sounds at encoding were motor corrected (see Method). The time windows defined for each peak were the following: Na 72–92 ms, Tb 120–150 ms, N1/N1<sub>mast</sub> 94–120 ms, P2/ P2<sub>mast</sub> 174–224 ms, P3 256–286 ms.

First, we performed a one-sided t test to test for possible differences in N1 amplitude between A and MA



**FIGURE 2** Summary of the behavioral results, separately for memory bias in the 2 T sequences (left) and memory performance in the 1 T sequences (right). Error bars depict standard error of the mean. Gray lines connect the data points of each subject, showing the response (% recall and % correct, respectively) to MA and A sounds for each individual. For memory bias (i.e., percent recall in 2 T sequences), there were no significant differences between motor-auditory and auditory-only sounds (two-tailed paired samples *t* test, *p* > .050, M<sub>*MA*</sub> = 0.509, M<sub>*A*</sub> = 0.491, SD<sub>*MA*</sub> = 0.0395, SD<sub>*A*</sub> = 0.0395), in line with the Bayesian analysis that provided weak evidence for the alternative hypothesis (*BF*<sub>10</sub> = 0.389). For memory performance (i.e., percent correct in 1 T sequences), there was a significant difference between motor-auditory and auditory-only sounds (two-tailed paired samples *t* test, *t*(22) = 3.15, *p* = .005, *d* = 0.66; indicated by two asterisks), with higher accuracy for the latter (M<sub>*MA*</sub> = 0.757, M<sub>*A*</sub> = 0.799, SD<sub>*MA*</sub> = 0.108, SD<sub>*A*</sub> = 0.0924), which was also supported by the Bayesian analysis that brought strong evidence in favor of the alternative hypothesis (*BF*<sub>10</sub> = 9.375).

sounds at encoding, with the hypothesis of attenuated responses for the latter. Indeed, we obtained a significant attenuation for the N1, t(22) = -1.89, p = .036, d = -0.39, with lower amplitudes for sounds that coincided with a motor act, compared to those that were passively presented to the participants (Figure 3a,b, see Table 1 for all the mean amplitudes per condition). We also tested for differences in N1 (with reversed polarity) at the mastoids (N1<sub>mast</sub>) using a repeated measures ANOVA with factors Sound Type (MA vs. A) and Laterality (M1 vs. M2). We obtained a significant enhancement for the MA sounds  $F(1, 22) = 15.68, p < .001, \eta_p^2 = .42$ , suggesting that besides the attenuation for MA sounds observed at vertex, further modulatory effects of sound-action coincidence occur (Figure 3). We also found a significant main effect of Laterality, F(1, 22) = 5.96, p = .023,  $\eta_p^2 = .21$ , with lower amplitudes at M1 compared to M2, while the interaction between Sound Type and Laterality did not reach significance, F(1, 22) = 3.55, p = .073.

Next, we examined the attenuation effects at the N1 subcomponents at temporal sites, using a 2×2 ANOVA, with factors Sound Type (A vs. MA) and Laterality (T7 vs. T8) on Na and Tb (Figure 3a). For Na, only a significant main effect of Laterality was obtained, with lower amplitudes at T8 compared to T7, F(1, 22) = 4.82, p = .039,  $\eta_p^2 = .18$ , while the main effect of Sound Type and the interaction did not reach significance, F(1, 22) = 0.05, p = .828 and F(1, 22) = 0.35, p = .563, respectively. For Tb, however, we obtained significantly lower amplitudes for sounds coinciding with a motor act compared to the auditoryonly ones, F(1, 22) = 9.03, p = .007,  $\eta_p^2 = .29$ , while the main effect of Laterality did not reach significance, F(1,22) = 0.03, p = .871. However, we also found a significant interaction, F(1, 22) = 8.63, p = .008,  $\eta_p^2 = .28$ , reflecting that the attenuation for MA sounds was only significant in T8 but not in T7 (post-hoc *t* tests, t(22) = -4.06, p < .001, d = -0.85 and t(22) = -1.04, p = .311, respectively).

Subsequently, we performed a one-sided *t* test to test for possible differences in P2 amplitudes between A and MA sounds at encoding, with the hypothesis of attenuated responses for the latter. We obtained a significant P2 attenuation at Cz, t(22) = 3.98, p < .001, d = 0.83, with lower amplitudes for sounds that coincided with a motor act, compared to those that were passively presented to the participants (Figure 3a,b). We also tested for differences in this component (with reversed polarity) at the mastoids (P2<sub>mast</sub>) using a repeated measures ANOVA with factors Sound Type (MA vs. A) and Laterality (M1 vs. M2). We observed a significant attenuation for the MA sounds, replicating the attenuation observed at Cz, F(1,22) = 34.23, p < .001,  $\eta_p^2 = .61$ , as well as a main effect of Laterality, F(1, 22) = 4.66, p = .042,  $\eta_p^2 = .17$ , with more negative amplitudes at M1 compared to M2. The interaction of Sound Type and Laterality on P2mast did not reach



**FIGURE 3** (a) Group-average event-related potentials across 23 participants for the corrected motor-auditory (red) and auditory-only (blue), analyzed in the corresponding electrodes. Difference waves (A–[MA–M]) depicting the self-generation effects are represented in black. Time windows used for the analyses are indicated in gray (Na: 72–92 ms, Tb: 120–150 ms, N1: 94–120 ms, P2: 174–224 ms, P3: 256–286 ms). Significant differences in the event-related potentials are indicated by asterisks. (b) N1, P2, and P3 scalp topographies in the time windows for: (1) the auditory-only condition (left); (2) the corrected motor-auditory condition (middle); and (3) the (A–[MA–M]) difference waves, reflecting suppression (N1, P2) and enhancement (P3) effects.

		Auditory-only (A)		Motor-auditory (MA)		Encoded as MA and forgotten		Encoded as MA and remembered		Encoded as A and forgotten		Encoded as A and remembered	
Components	Electrodes	Μ	SD	М	SD	Μ	SD	Μ	SD	Μ	SD	Μ	SD
N1	Cz	-3.14	1.79	-2.66	1.98	-3.89	2.01	-4.51	2.28	-4.13	2.45	-4.19	2.22
P2	Cz	4.95	2.49	3.83	2.01	7.16	4.38	7.37	3.51	7.33	3.96	7.76	4.25
P3	Pz	-0.08	1.29	1.49	1.43	-	-	-	-	-	-	-	-
N1 <sub>mast</sub>	M1	0.26	0.87	0.67	0.84	0.51	1.33	0.27	1.02	0.59	0.95	0.53	1.29
	M2	0.43	0.99	1.03	0.98	0.65	1.12	0.61	1.41	0.83	1.38	0.86	1.33
P2 <sub>mast</sub>	M1	-0.75	0.79	-0.19	0.81	-1.88	1.51	-2.53	1.71	-2.03	1.32	-2.24	1.42
	M2	-0.56	1.01	0.05	0.87	-2.24	1.43	-2.63	1.64	-2.18	1.57	-2.45	1.55
Na	T7	-0.89	0.94	-0.97	1.18	-1.23	1.37	-1.48	1.19	-1.11	1.02	-0.86	1.09
	T8	-0.47	0.76	-0.45	1.03	-0.89	1.36	-1.21	1.30	-0.82	1.68	-0.59	1.12
Tb	T7	-1.91	1.01	-1.75	1.12	-2.89	1.73	-3.26	1.94	-2.97	1.66	-2.34	1.53
	T8	-2.18	1.40	-1.54	1.56	-3.68	2.25	-3.62	1.94	-3.40	2.19	-2.81	1.63

TABLE 1 Mean amplitudes and standard deviation per component and condition across 23 participants

significance, F(1, 22) = 0.54, p = .470. Finally, we also tested for differences in P3 at Pz, which yielded a significantly larger P3 amplitude for sounds coinciding with a motor act, t(22) = -6.57, p < .001, d = -1.37 (Figure 3). Finally, we decided to examine our data using a more datadriven approach to test for further effects that may have not been captured in the hypotheses-driven ERP analysis (cluster-based permutation analyses; see Supporting Information; Maris & Oostenveld, 2007). We found a negative cluster (p < .001; 56–344 ms post-stimulus) and one positive cluster (p = .01; 122–232 ms post-stimulus), in line with the findings obtained in the ERP analysis (see Supporting Information).

# 3.3 | Electrophysiological responses at retrieval

Next, we performed exploratory analyses for the retrieval data, by subdividing it depending on whether the sound was encoded as A or MA and whether this sound was recalled or not. This allowed us to assess whether auditory evoked responses were affected by how the sound was encoded and whether it was remembered or forgotten. To this end, we ran an ANOVA with Sound Type (Encoded as MA vs. Encoded as A) and Memory (Remembered vs. Forgotten) as within-subject factors on N1/N1<sub>mast</sub>, P2/ P2<sub>mast</sub>, Na, and Tb. An electrode factor (Laterality) was included in the ANOVA for the components identified in the mastoids and temporal electrodes. Figure 4 shows all the studied peaks for the remembered (a) and the forgotten (b) sounds at retrieval in the time windows 72-92, 120-150, 94–120, 174–224 ms, for the Na, Tb, N1/N1<sub>mast</sub>, and P2/ P2<sub>mast</sub>, respectively at the relevant electrodes for each peak.

We did not observe any significant effects (all ps > .05) on the N1 at Cz and N1<sub>mast</sub>. However, significant results were obtained when we analyzed the modulatory effects of Sound Type and Memory on the N1 subcomponents at temporal sites. We obtained a significant main effect of Sound Type on Na, F(1, 22) = 7.39, p = .013,  $\eta_p^2 = .25$ , and Tb, F(1, 22) = 7.28, p = .013,  $\eta_p^2 = .25$ , reflecting an enhanced amplitude for sounds that were previously encoded as MA. Additionally, we found a significant interaction between Sound Type and Memory on Na, F(1,22) = 5.08, p = .035,  $\eta_p^2 = .19$ , where post-hoc comparisons showed significantly larger Na amplitude for sounds that were Encoded as MA and were remembered compared to sounds that were Encoded as A and were remembered, t(45) = 3.73, p < .001, d = 0.55. In contrast, the post-hoc comparisons did not show significant differences for forgotten sounds as a function of how they were encoded, t(45) = 0.67, p = .504. No significant differences were found between remembered and forgotten sounds that

# PSYCHOPHYSIOLOGY SPR

were Encoded as A, t(45) = -1.34, p = .187, or between remembered and forgotten sounds that were Encoded as MA, t(45) = 1.64, p = .109. Similarly, we obtained a significant interaction between Sound Type and Memory on Tb,  $F(1, 22) = 4.85, p = .038, \eta_p^2 = .18$ . Post-hoc comparisons showed significantly larger Tb amplitude for sounds that were Encoded as MA and were remembered compared to sounds that were Encoded as A and were remembered, t(45) = 4.31, p < .001, d = 0.64, which is in line with the differences we obtained in the Na window. The posthoc comparisons also showed lower Tb amplitudes for the Encoded as A sounds when they were remembered compared to when they were forgotten, t(45) = -3.23, p = .002, d = -0.48. Nevertheless, no significant differences were observed between remembered and forgotten sounds that were encoded as MA, t(45) = 0.64, p = .523, or between the Encoded as MA and Encoded as A sounds that were forgotten, t(45) = 0.47, p = .640. For both Na and Tb, we did not observe any significant main effects of Laterality, nor any significant interactions between Laterality and Sound Type and/or Memory (all ps > .05). Finally, we did not observe any significant effects on P2 at Cz and P2<sub>mast</sub> (all ps > .05), except for a significant main effect of Memory on  $P2_{mast}$ , F(1, 22) = 7.65, p = .011,  $\eta_p^2 = .26$ , that showed lower amplitudes for sounds that were forgotten ( $M_{Forgotten} = -2.08$ ,  $M_{Remembered} = -2.46$ ,  $SD_{Forgotten} = 1.44$ ,  $SD_{Remembered} = 1.56$ ). Similar to the approach we followed for the encoding data, we also conducted exploratory analyses using cluster-based permutation statistics (Maris & Oostenveld, 2007), but we did not find any significant clusters for any of the effects (see Supporting Information).

# 3.4 | Pupil responses at encoding and retrieval

Cluster-based permutation statistics were used to test for possible differences in pupil diameter between the conditions of interest. First, we tested for differences in the pupil response between motor-auditory and auditory-only events at encoding and we obtained significantly larger pupil diameter for motor-auditory events (starting 180 ms before sound onset and lasting up to 1230 ms after sound onset; p < .05; Figure 5a) in line with previous work in rodents (e.g., McGinley et al., 2015). Interestingly, the effect of action started already in the pre-stimulus period, that is before the button press (which immediately triggered the sound), in agreement with previous work showing that LC activity and pupil diameter start increasing before the onset of movement (Aston-Jones & Cohen, 2005; Reimer et al., 2016). Subsequently, we conducted an exploratory analysis to test for possible main effects of Sound







**FIGURE 4** Group-average event-related potentials across 23 participants for the encoded as MA (red) and encoded as A (blue), analyzed in the corresponding electrodes and presented separately for the remembered (left) and the forgotten sounds (right). Time windows used for the analyses are indicated in gray (Na: 72–92 ms, Tb: 120–150 ms, N1: 94–120 ms, P2: 174–224 ms). Significant differences in the event-related potentials are indicated by asterisks.

Type (Encoded as A vs. Encoded as MA) and Memory (Remembered vs. Forgotten), as well as for interactions between Sound Type and Memory on the pupil responses at retrieval. This analysis showed only a significant main effect of Memory, with larger diameter for forgotten sounds at retrieval compared to the remembered ones, irrespective of how they were encoded (starting 170 ms after sound onset and lasting until 830 ms after sound onset; p < .05; Figure 5b). Note that the morphology of the responses differs between the encoding (Figure 5a) and the retrieval (Figure 5b) data, most likely due to differences

in the visual stimulation between the two phases (i.e., dynamic visual stimulation with the moving line at encoding vs. brief and static visual stimuli at retrieval, namely the cues "Sound 1" and "Sound 2").

# 3.5 | Correlations

Next, we tested for possible correlations between the behavioral performance, pupillometric and electrophysiological data. For the correlation analyses, we focused

14 of 25



**FIGURE 5** (a) The group-average evoked pupil responses at encoding to auditory-only (blue) and motor-auditory (red) events. The effect is depicted as the difference between auditoryonly and motor-auditory events (black). Black bar indicates a significant auditory-only vs. motor-auditory effect in the window 180 pre-stimulus to 1230 ms post-stimulus, p < .05 (cluster-based permutation test). (b) The group-average evoked pupil responses at retrieval to encoded as auditory (A) and encoded as motor-auditory (MA), separately for the remembered and forgotten sounds. Black bar indicates a significant main effect of memory for remembered vs. forgotten sounds in the window 170–830 ms post-stimulus, p < .05 (cluster-based permutation test).

on the significant neurophysiological effects at encoding (i.e., ERPs and pupil diameter) and the significant behavioral effect on memory performance. The effects were introduced in the correlation analyses as the difference between A and MA events (see Method). For the components identified in two electrodes, we calculated the mean amplitude across the two, except for the Tb at encoding, where we introduced only the amplitudes at T8 given the significant interaction between Sound Type and Laterality that showed that attenuation was lateralized. For the pupil data, we calculated the peak of the difference wave (A – MA) within the window of significance (180 ms prestimulus until 1230 ms post-stimulus). All the planned correlations are reported in Table 2. PSYCHOPHYSIOLOGY SPR

15 of 25

4698986, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14156 by Readcube (Labtiva Inc.), Wiley Online Library on [28/10/2024]. See the Terms and Conditions

(https://onlinelibrary.wiley.com

and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 2** Correlations between the significant self-generation effects. (a) Electrophysiological effects at encoding (N1, P2,  $N1_{mast}$ , P2<sub>mast</sub>, P3, and Tb amplitudes) and memory performance (1 T sequences), (b) neuromodulatory effects at encoding (pupil diameter) and memory performance (1 T sequences), (c) electrophysiological (N1, P2,  $N1_{mast}$ , P2<sub>mast</sub>, P3, and Tb amplitudes) and neuromodulatory (pupil diameter) effects at encoding.

Correlations between	r	р				
(a) Memory performance (1 T sequences)						
N1	43	.041*				
Tb (at T8 only)	55	.007**				
P2	19	.383				
N1 <sub>mast</sub>	41	.055				
P2 <sub>mast</sub>	10	.657				
P3	35	.098				
(b) Memory performance (1 T s	equences)					
Pupil diameter	.46	.029*				
(c) Pupil diameter						
N1	36	.091				
Tb (at T8 only)	25	.251				
P2	.27	.209				
N1 <sub>mast</sub>	23	.291				
P2 <sub>mast</sub>	16	.507				
Р3	08	.702				

Significant correlations are highlighted in bold and indicated by asterisks (\*p < .05 and \*\*p < .01).

First, we tested whether the significant self-generation effects at encoding (on N1, P2, N1<sub>mast</sub>, P2<sub>mast</sub>, P3, and Tb amplitudes) correlated with the significant self-generation effects on memory performance (1 T sequences). This analysis showed a negative correlation between N1 suppression and memory performance (r = -.43, p = .041; Figure 6a), and a negative correlation between Tb suppression (at T8) and memory performance (r = -.55, p = .007; Figure 6b), that is, the larger the N1 and Tb suppression, the greater the memory impairment for motor-auditory compared to auditory-only sounds. The remaining correlations did not reach significance (all ps > .05). Second, we assessed whether the difference in pupil diameter between auditory-only and motor-auditory events was related to memory performance and we obtained a significant positive correlation between the two (r = .46, p = .029; Figure 6c), that is, the larger the pupil dilation for the motor-auditory events, the greater the memory impairment for these sounds. Third, we tested for possible links between the self-generation effects obtained in the ERP analyses (i.e., N1, P2, N1<sub>mast</sub>, P2<sub>mast</sub>, P3 and Tb) and the larger pupil diameter for motor-auditory events. None of these correlations reached significance (all ps > .05), but we observed a non-significant trend toward a correlation



**FIGURE 6** Planned correlations between the behavioral, electrophysiological, and pupil data using the Pearson correlation coefficient. (a and b) Significant negative correlations between N1 suppression (at Cz) and memory performance (r = -.43, p = .041), and Tb suppression (at T8) and memory performance (r = -.55, p = .007), showing that the larger the N1 and Tb suppression, the greater the memory impairment for motor-auditory compared to the auditory-only sounds. More negative values indicate larger suppression effects for N1 and Tb and worse memory performance for motor-auditory sounds. (c) Significant positive correlation between pupil dilation and memory performance (r = .46, p = .029), that is, the larger the pupil dilation for the motor-auditory events, the greater the memory impairment for these sounds. (d) The correlation between N1 attenuation at Cz and pupil dilation at encoding for the MA events did not reach significance (r = -.36, p = .091). The shaded gray areas represent the confidence interval (95% confidence level).

between N1 attenuation at Cz and pupil dilation for MA events (Figure 6d).

Finally, we performed an exploratory correlation analysis to test whether the significant differences in sensory processing we obtained at retrieval between Encoded as A and Encoded as MA sounds were related to the magnitude of the self-generation effects at encoding. To this end, we performed a correlation analysis between the A - MA difference in peaks of the Na and Tb amplitudes (only for the remembered sounds due to the significant interaction) and the effects at encoding (for the N1, P2, N1<sub>mast</sub>, P2<sub>mast</sub>, P3, and Tb amplitudes). We obtained a significant positive correlation between the P2 suppression at encoding and the Na enhancement at retrieval for the remembered sounds, reflecting that the larger the attenuation for P2 at encoding, the larger the Na enhancement for the Encoded as MA sounds that were remembered at retrieval (r = .51, p = .012). Similarly, we also obtained a significant negative correlation between Tb at encoding (at T8) and Na for the remembered sounds at retrieval (r = -.42, p = .04),

ing, the greater the Na enhancement for motor-auditory sounds that were remembered at retrieval.

showing that the larger the attenuation for Tb at encod-

# 4 | DISCUSSION

In this study, we assessed the effects of motor actions on sensory processing and memory encoding of concomitant, but unpredictable sounds, by employing a combination of a self-generation and memory recognition task, while monitoring the brain's and the pupil's responses to sounds that were either presented passively or that coincided in time with a motor act. The aim of the present work was to assess how motor acts affect first sensory processing and second memory encoding of concomitant sounds, and the possible relationships between these two types of effects of actions. Related to the first aim, regarding the effects of actions on sensory processing, we examined whether (a) attenuation of sensory processing (i.e., measured by ERPs) prevails even in the absence of a contingent action-sound

16 of 25

relationship (e.g., Horváth et al., 2012), (b) actions create a halo of subcortical neuromodulation around them that could be reflected in the pupil diameter (e.g., McGinley et al., 2015), and (c) sensory processing (i.e., measured by ERPs) and subcortical neuromodulation (i.e., measured by pupil diameter) during actions were related. Our findings showed N1, P2, P2<sub>mast</sub>, and Tb attenuation for motor-auditory sounds even when they merely coincide with the action, as well as enhancement of P3 and N1<sub>mast</sub> (cf. Horváth et al., 2012). These findings suggest that selfgeneration effects are at least partly stimulus-unspecific and driven by alternative mechanisms to the cancellation of predicted sensory reafference via motor forward modeling. Additionally, our data replicated previous work (e.g., Lee & Margolis, 2016; McGinley et al., 2015; Simpson, 1969; Vinck et al., 2015; Yebra et al., 2019) showing that pupil diameter increases dramatically during actions providing evidence for an alternative stimulusunspecific mechanism that could partly underlie sensory suppression for self-generated sounds, namely the activation of subcortical neuromodulation during motor actions. However, contrary to our initial hypothesis, the data did not provide clear evidence for a correlation between sensory attenuation and pupil dilation for motor-auditory events. The second aim of the present study was to investigate how actions affect memory encoding of concomitant sounds and whether the potential differences in the memory encoding of motor-auditory and passively presented sounds correlate with sensory suppression and/or subcortical neuromodulation during encoding. We found a significant impairment in memory performance for sounds that were encoded as motor-auditory compared to the auditory-only ones demonstrating that the mere presence of an action affects memory encoding of simultaneously presented stimuli. Most importantly, worsened memory performance for motor-auditory events correlated with increased sensory suppression (i.e., N1 and Tb attenuation) and larger pupil dilation for motor-auditory events at encoding. These findings fit well with the predictive coding framework suggesting that prediction errors (i.e., reflected in ERPs) drive learning and memory (Henson & Gagnepain, 2010) and further support previous work showing that high arousal (i.e., reflected in pupil diameter) may worsen behavioral performance (McGinley et al., 2015). In the following, we discuss each of these effects in detail.

The first aim of the present study was to assess the effects of actions on auditory processing and subcortical neuromodulation, as well as the relationship between the two. First, we provide evidence that the self-generation effects are at least partly unspecific by showing that N1 attenuation prevails even for mere action-sound coincidences and that it partly reflects the modulation of the

PSYCHOPHYSIOLOGY SPR

4698986, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14156 by Readcute (Labiva Inc.), Wiley Online Library on [28/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/term

and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

unspecific N1 component, as for the suppression to be specific to the auditory cortex, N1 should be suppressed at vertex but also at the mastoids, which was not found here (cf. Horváth, 2013b; Horváth et al., 2012). This finding goes along with previous work pointing to partly unspecific mechanisms behind the action-induced suppression effects (e.g., Horváth et al., 2012; SanMiguel et al., 2013). For example, attenuation of auditory responses occurs also for stimuli merely coinciding with finger movements (Hazemann et al., 1975; Horváth et al., 2012; Makeig et al., 1996; Tapia et al., 1987) or for unrelated auditory inputs during speech (Numminen et al., 1999). Similarly, previous work has suggested that N1 (and Tb) attenuation can be driven by mere temporal contiguity (Han et al., 2021; Hazemann et al., 1975; Horváth et al., 2012) or by temporal predictability (Kaiser & Schütz-Bosbach, 2018; Lubinus et al., 2021; Schafer & Marcus, 1973, but see also Klaffehn et al., 2019 for evidence showing that attenuation prevails when controlling for temporal predictions), rather than stimulus-specific predictions, and that it mostly reflects modulations of the unspecific component of the auditory N1 (SanMiguel et al., 2013). Meanwhile, there is also mounting evidence supporting the stimulusspecificity of the effects by showing more pronounced suppression when predictions match more precisely with the sensory input (Baess et al., 2008; Fu et al., 2006; Hashimoto & Sakai, 2003; Heinks-Maldonado et al., 2005; Houde et al., 2002). Collectively, we believe that our findings point to the involvement of unspecific processes in the action-induced suppression of auditory responses that can, nevertheless, co-exist with stimulus-specific predictive mechanisms as suggested by previous work (Flinker et al., 2010; Horváth, 2015; Schröger et al., 2015).

In addition to the N1-attenuation effects, we observed attenuated P2 and enhanced P3 responses for the sounds coinciding with actions. Although a functional interpretation of P2 is missing (Crowley & Colrain, 2004), empirical evidence has shown that the P2 component originates in secondary auditory areas (Bosnyak et al., 2004; Pantev et al., 1996), reflecting the processing of the specific features of auditory stimuli (Shahin et al., 2005), and it correlates with the sense of agency (i.e., the feeling of control over actions and their consequences; Gallagher, 2000) contrary to the N1 that does not (Ford et al., 2014; Kühn et al., 2011; Timm et al., 2016). These characteristics along with our data showing P2 attenuation in both vertex and mastoids may point to a functional dissociation between N1 and P2 as suggested by previous work (Chen et al., 2012; Knolle et al., 2013b; Schröger et al., 2015). Following the P2 attenuation, we found enhanced P3 amplitude at Pz for sounds coinciding with actions. Interestingly, a P3 effect was also evident - although not discussed - in previous work with action-sound coincidences (Horváth et al., 2012).

Recently, this effect has been suggested to reflect violations in action-related predictions (Darriba et al., 2021) which may occur in tasks where the self-generated sound is unexpected (e.g., in coincidence tasks where the action does not always result in a sound; Horváth et al., 2012). Although previous work has already described P3 modulations in self-generation paradigms, the posterior distribution and later peak of our effect differentiates it from the fronto-central P3a effect reported for unexpected externally-generated sounds (Baess et al., 2011) or selfgenerated deviant sounds (Knolle et al., 2013a). Based on previous theories, we speculate that the posterior P3 effect may be related to context updating (Donchin & Coles, 1988), event categorization (Kok, 2001) or decision making (Twomey et al., 2015) and may reflect an evaluative process of the stimulus (i.e., self/external categorization) that ultimately updates the internal model about the sensory consequences of the button press (Polich, 2007).

The second important finding related to our first aim is that neuromodulatory processes take place concomitantly to the modulatory effects of action-sound coincidence on evoked electrophysiological responses. We obtained pupil dilation measures that are known to track the activity of the LC-NE system (Aston-Jones & Cohen, 2005; Joshi et al., 2016; Murphy et al., 2014) and in line with our hypothesis, we showed a remarkable increase in pupil diameter for the motor-auditory events that started even before the action (cf. Aston-Jones & Cohen, 2005; McGinley et al., 2015; Reimer et al., 2016), supporting previous work reporting pupil dilation during finger movements (Lubinus et al., 2021; Yebra et al., 2019), and locomotion (McGinley et al., 2015; Reimer et al., 2014; Vinck et al., 2015) even in the absence of visual stimulation (Hupe et al., 2009). We also hypothesized that these neuromodulatory processes might be behind the stimulus-unspecific effects of actions on the auditory evoked responses. However, pupil dilation did not correlate with the sensory suppression effects for self-generated sounds. Although this may suggest that motor-induced sensory suppression and arousal-related neuromodulation during actions operate independently, there was a non-significant trend toward a link between N1 attenuation at vertex and pupil dilation, and both of these measures correlated significantly with memory performance. Taken together, these findings raise the need of future work to further test for relationships between action-induced suppression effects and neuromodulatory mechanisms operating during movement.

The second aim of the present study was to assess how the differential processing for sounds coinciding with actions might affect their encoding in memory. While the links between sensorimotor processing of auditory stimuli and memory processes remain largely unexplored, there is evidence that actions attenuate responses in areas supporting memory processes (i.e., Mukamel et al., 2010; Rummell et al., 2016), raising the possibility of a link between self-generation and memory. In our study, motor actions affected the memory encoding of concurrent sounds, but the effects were reflected only in memory performance and not in memory bias. The null effects on memory bias might suggest that participants could recognize that both test sounds at retrieval were presented before, which is supported by the general high level of objective accuracy as well as by reports during an informal debriefing suggesting that many participants thought that most times all sounds at retrieval were presented before. The memory benefit for the more surprising externally generated sounds fits well with predictive coding theories postulating that items eliciting larger prediction errors at encoding will be encoded better in memory (Exton-McGuinness et al., 2015; Greve et al., 2017; Heilbron & Chait, 2018; Henson & Gagnepain, 2010; Krawczyk et al., 2017; Pine et al., 2018; Rescorla & Wagner, 1972). Yet, one would expect to observe this effect only in contingent paradigms where self-generated sounds are inherently more predictable than the externally generated ones. However, although in our study actions were not predictive of sound identity or occurrence, they afforded better temporal predictability, which might have rendered motor-auditory sounds less salient, thereby compromising their encoding in memory (but not in 2 T sequences where participants clearly remembered both sounds). We, therefore, acknowledge that our study cannot completely disentangle whether the effects observed on memory encoding are due to the neurophysiological effects of motor acts at encoding (e.g., attenuation and increased neuromodulation as indexed by pupil dilation), temporal predictability, or both.

Related to the second aim of the present study, we also hypothesized that the memory encoding of sounds paired with actions should be related to the neurophysiological effects of actions on sensory processing of sounds, namely the suppression effects and the pupil dilation for action-sound coincidences. First, we showed that the self-generation effects (i.e., N1 and Tb attenuation) are related to the performance decrements for sounds produced by actions as suggested by previous work in rodents (McGinley et al., 2015; Schneider et al., 2014, for a review see Schneider, 2020). These findings support the idea that the larger prediction error responses to unexpected items (as indexed by enhanced ERPs to A compared to MA events at encoding) initiate a cascade of synaptic changes, allowing for more distinctive representations at encoding (Kirwan & Stark, 2007; Norman, 2010) and thus better recollection at retrieval. Our findings could also fit with the compelling evidence for hippocampal involvement in learning from prediction errors (Schiffer et al., 2012) and

expecting upcoming events (Davachi & DuBrow, 2015; Hindy et al., 2016; Schapiro et al., 2017): The reduced prediction errors at hippocampus to self-initiated stimulation (Mukamel et al., 2010; Rummell et al., 2016) could translate to memory decrements for these items. Second, we showed that memory performance correlated with pupil diameter as well, such that the larger the pupil diameter for motor-auditory events the worse the memory performance for these sounds at retrieval. To date, there have been no direct attempts to test for possible links between motor-induced pupil dilation and memory performance for stimuli triggered by actions. Some interim evidence points to a negative relationship between pupil dilation and detection performance during locomotion (McGinley et al., 2015), suggesting that performance may follow the classically described, inverted U-shaped dependence on arousal (Yerkes & Dodson, 1908): Intermediate levels of arousal – as indexed by pupil diameter – occur in states of quiet wakefulness and are characterized by optimal performance. In contrast, performance during high-arousal states such as movement drops dramatically. Collectively, we showed that sensory attenuation and pupil dilation independently correlate with memory performance, supporting the predictive account of memory (i.e., memory enhancements for items eliciting larger prediction errors at encoding) and providing yet another piece of evidence supporting the detrimental effects of high arousal (i.e., as indexed by pupil diameter) on behavioral performance.

The present study had clear hypotheses about the effects of actions on sensory and pupil responses at encoding, yet, exploratory analyses of the retrieval data revealed further effects. First, we obtained higher Na and Tb amplitudes for the sounds encoded as motor-auditory and remembered compared to the remembered and encoded as auditory-only ones. As the sounds encoded as motor-auditory were presented passively at retrieval (i.e., without the motor representation that they were encoded with), the higher Na and Tb amplitudes may reflect a form of contextual prediction error (Exton-McGuinness et al., 2015; Kim et al., 2014; Sinclair & Barense, 2019) due to the mismatch between encoding and retrieval contexts for these sounds. This interpretation can be partly supported by the exploratory correlation analyses that showed that the larger the P2 and Tb attenuation for motor-auditory sounds at encoding, the greater the Na enhancement for these sounds at retrieval when they were remembered. Thus, the greater the effect of the action at encoding, the greater the contextual prediction error when the sound is presented without the action at retrieval. Second, we found larger pupil responses for the forgotten compared to the remembered sounds at retrieval irrespective of how they were encoded. While previous work has reported an old/new pupil effect (i.e., increased pupil responses for the remembered items; Kafkas & Montaldi, 2015; Naber

PSYCHOPHYSIOLOGY SPR

et al., 2013, but see Beukema et al., 2019 for the opposite effect), in our study both sounds at retrieval were presented before. The increase in pupil diameter for the forgotten sounds at retrieval could be instead related to selection or decision uncertainty (Geng et al., 2015; Nassar et al., 2012; Preuschoff et al., 2011; Richer & Beatty, 1987) when participants experienced greater difficulty to decide whether a given sound was presented before or not.

In sum, the overarching aim of the present study was to investigate how motor acts affect both sensory processing and the memory encoding of concomitant sounds. To the best of our knowledge, there have been no previous attempts to simultaneously assess the specificity of the self-generation effects and their possible link with neuromodulatory processes while also looking into the effects of actions on memory encoding of sounds. Here, in a combination of self-generation and memory task, we show that actions affect auditory responses, pupil diameter, and memory encoding of sounds. Actions suppressed sensory responses for concomitant sounds and increased pupil diameter, but these effects were not related, pointing to simultaneous, but probably independent processes. However, sensory suppression and pupil dilation both correlated with memory performance independently, such that the memory performance for sounds coinciding with actions decreased with larger sensory attenuation and greater pupil dilation. Collectively, our findings show self-generation effects even in the absence of a predictive action-sound relationship, replicate previous work showing that pupil diameter increases during actions, and finally point to differentiated internal memory representations for stimuli triggered by ourselves compared to externally presented ones. More importantly, the present study shows that subcortical neuromodulatory systems, along with cortical processes, simultaneously orchestrate auditory processing and memory encoding.

#### AUTHOR CONTRIBUTIONS

**Nadia Paraskevoudi:** Conceptualization; formal analysis; investigation; methodology; software; writing – original draft. **Iria SanMiguel:** Conceptualization; formal analysis; funding acquisition; methodology; project administration; software; supervision; writing – original draft.

#### ACKNOWLEDGMENTS

This work is part of the project PSI2017-85600-P, funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe"; it has been additionally supported by the MDM-2017-0729-18-2M Maria de Maeztu Center of Excellence UBNeuro, funded by MCIN/AEI/ 10.13039/501100011033, and by the Excellence Research Group 2017SGR-974 funded by the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la

Generalitat de Catalunya. ISM was supported by grant RYC-2013-12577, funded by MCIN/AEI/10.13039/501100011033 and by "ESF Investing in your future." NP was supported by predoctoral fellowship FI-DGR 2019 funded by the Secretaria d'Universitats i Recerca de la Generalitat de Catalunya and the European Social Fund.

## **CONFLICT OF INTEREST**

None.

## ORCID

Nadia Paraskevoudi https://orcid. org/0000-0001-9946-7580 Iria SanMiguel https://orcid.org/0000-0002-7462-8621

#### REFERENCES

- Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience*, 21(4), 791–802. https://doi.org/10.1162/jocn.2009.21055
- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, 28(1), 403– 450. https://doi.org/10.1146/annurev.neuro.28.061604.135709
- 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
- Baess, 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
- Baess, P., Widmann, A., Roye, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *European Journal of Neuroscience*, 29(7), 1514–1521. https://doi. org/10.1111/j.1460-9568.2009.06683.x
- Bar, M. (2009). The proactive brain: Memory for predictions. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1521), 1235–1243. https://doi.org/10.1098/rstb.2008.0310
- 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/journ al.pbio.0040028
- Beukema, S., Jennings, B. J., Olson, J. A., & Kingdom, F. A. A. (2019). The pupillary response to the unknown: Novelty versus familiarity. *I-Perception*, 10(5), 204166951987481. https://doi. org/10.1177/2041669519874817
- Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal* of Neuroscience, 33(5), 2199–2204. https://doi.org/10.1523/ JNEUROSCI.3440-12.2013
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7), 635–640. https://doi.org/10.1038/2870
- Bornert, P., & Bouret, S. (2021). Locus coeruleus neurons encode the subjective difficulty of triggering and executing actions.

*PLoS Biology*, *19*(12), e3001487. https://doi.org/10.1371/journ al.pbio.3001487

- Bosnyak, D. J., Eaton, R. A., & Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex*, 14(10), 1088–1099. https://doi. org/10.1093/cercor/bhh068
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357
- Brown, R. M., & Palmer, C. (2012). Auditory–motor learning influences auditory memory for music. *Memory & Cognition*, 40(4), 567–578. https://doi.org/10.3758/s13421-011-0177-x
- Chagnaud, B. P., Banchi, R., Simmers, J., & Straka, H. (2015). Spinal corollary discharge modulates motion sensing during vertebrate locomotion. *Nature Communications*, 6(1), 7982. https:// doi.org/10.1038/ncomms8982
- Chapman, C. E., Bushnell, M. C., Miron, D., Duncan, G. H., & Lund, J. P. (1987). Sensory perception during movement in man. *Experimental Brain Research*, 68(3), 516–524. https://doi. org/10.1007/BF00249795
- Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. *American Journal of EEG Technology*, 25(2), 83–92. https://doi.org/10.1080/00029238.1985.11080163
- Chen, Z., Chen, X., Liu, P., Huang, D., & Liu, H. (2012). Effect of temporal predictability on the neural processing of self-triggered auditory stimulation during vocalization. *BMC Neuroscience*, *13*(1), 55. https://doi.org/10.1186/1471-2202-13-55
- Cohen, L. G., & Starr, A. (1987). Localization, timing and specificity of gating of somatosensory evoked potentials during active movement in man. *Brain*, *110*(2), 451–467. https://doi. org/10.1093/brain/110.2.451
- Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, *26*(3), 341–361. https://doi.org/10.1016/0749-596X(87)90118-5
- 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
- Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, 115(4), 732–744. https:// doi.org/10.1016/j.clinph.2003.11.021
- Darriba, Á., Hsu, Y.-F., Van Ommen, S., & Waszak, F. (2021). Intentionbased and sensory-based predictions. *Scientific Reports*, *11*(1), 19899. https://doi.org/10.1038/s41598-021-99445-z
- Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99. https://doi.org/10.1016/j. tics.2014.12.004
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–374. https://doi.org/10.1017/S0140525X00058027
- Eggermann, E., Kremer, Y., Crochet, S., & Petersen, C. C. H. (2014). Cholinergic signals in mouse barrel cortex during active whisker sensing. *Cell Reports*, 9(5), 1654–1660. https://doi. org/10.1016/j.celrep.2014.11.005

- Eliades, S. J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology*, 89(4), 2194–2207. https://doi. org/10.1152/jn.00627.2002
- Exton-McGuinness, M. T. J., Lee, J. L. C., & Reichelt, A. C. (2015). Updating memories—The role of prediction errors in memory reconsolidation. *Behavioural Brain Research*, 278, 375–384. https://doi.org/10.1016/j.bbr.2014.10.011
- Flinker, A., Chang, E. F., Kirsch, H. E., Barbaro, N. M., Crone, N. E., & Knight, R. T. (2010). Single-trial speech suppression of auditory cortex activity in humans. *Journal of Neuroscience*, 30(49), 16643–16650. https://doi.org/10.1523/JNEUR OSCI.1809-10.2010
- 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
- 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
- Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions* of the Royal Society of London. Series B: Biological Sciences, 355(1404), 1771–1788. https://doi.org/10.1098/rstb.2000.0734
- Fu, C. H. Y., Vythelingum, G. N., Brammer, M. J., Williams, S. C. R., Amaro, E., Andrew, C. M., Yágüez, L., van Haren, N. E. M., Matsumoto, K., & McGuire, P. K. (2006). An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. *Cerebral Cortex*, 16(7), 969–977. https://doi.org/10.1093/ cercor/bhj039
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14–21. https://doi.org/10.1016/S1364-6613(99)01417-5
- Geng, J. J., Blumenfeld, Z., Tyson, T. L., & Minzenberg, M. J. (2015). Pupil diameter reflects uncertainty in attentional selection during visual search. *Frontiers in Human Neuroscience*, 9, 435. https://doi.org/10.3389/fnhum.2015.00435
- Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, 94, 149–165. https://doi. org/10.1016/j.jml.2016.11.001
- Halgren, E. (1991). Firing of human hippocampal units in relation to voluntary movements. *Hippocampus*, 1(2), 153–161. https:// doi.org/10.1002/hipo.450010204
- Han, N., Jack, B. N., Hughes, G., Elijah, R. B., & Whitford, T. J. (2021). Sensory attenuation in the absence of movement: Differentiating motor action from sense of agency. *Cortex*, 141, 436–448. https://doi.org/10.1016/j.cortex.2021.04.010
- Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Human Brain Mapping*, 20(1), 22–28. https://doi.org/10.1002/hbm.10119
- Hazemann, P., Audin, G., & Lille, F. (1975). Effect of voluntary self-paced movements upon auditory and somatosensory evoked potentials in man. *Electroencephalography* and Clinical Neurophysiology, 39(3), 247–254. https://doi. org/10.1016/0013-4694(75)90146-7
- 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., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, *20*(11), 1315–1326. https:// doi.org/10.1002/hipo.20857
- Hesse, M. D., Nishitani, N., Fink, G. R., Jousmaki, V., & Hari, R. (2010). Attenuation of somatosensory responses to selfproduced tactile stimulation. *Cerebral Cortex*, 20(2), 425–432. https://doi.org/10.1093/cercor/bhp110
- Hindy, N. C., Ng, F. Y., & Turk-Browne, N. B. (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature Neuroscience*, 19(5), 665–667. https://doi. org/10.1038/nn.4284
- Hoeks, B., & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments, & Computers, 25*(1), 16–26. https://doi. org/10.3758/BF03204445
- Horváth, J. (2013a). Attenuation of auditory ERPs to action-sound coincidences is not explained by voluntary allocation of attention. *Psychophysiology*, 50(3), 266–273. https://doi.org/10.1111/ psyp.12009
- Horváth, J. (2013b). Action-sound coincidence-related attenuation of auditory ERPs is not modulated by affordance compatibility. *Biological Psychology*, *93*(1), 81–87. https://doi.org/10.1016/j. biopsycho.2012.12.008
- Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54–65. https://doi.org/10.1016/j.brainres.2015.03.038
- Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action–sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience*, 24(9), 1919–1931. https://doi.org/10.1162/jocn\_a\_00215
- Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience*, 14(8), 1125– 1138. https://doi.org/10.1162/089892902760807140
- Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, 139(1), 133–151. https://doi. org/10.1037/a0028566
- Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual sensory attenuation in voluntary action. *NeuroImage*, 56(3), 1632–1640. https://doi.org/10.1016/j.neuro image.2011.02.057
- Hupe, J. M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision*, 9(7), 10. https://doi.org/10.1167/9.7.10
- Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*(1), 221–234. https://doi.org/10.1016/j.neuron.2015.11.028
- Kafkas, A., & Montaldi, D. (2015). The pupillary response discriminates between subjective and objective familiarity and novelty. *Psychophysiology*, 52(10), 1305–1316. https://doi.org/10.1111/ psyp.12471
- 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

- Kelley, D. B., & Bass, A. H. (2010). Neurobiology of vocal communication: Mechanisms for sensorimotor integration and vocal patterning. *Current Opinion in Neurobiology*, 20(6), 748–753. https://doi.org/10.1016/j.conb.2010.08.007
- Kilteni, K., Engeler, P., & Ehrsson, H. H. (2020). Efference copy is necessary for the attenuation of self-generated touch. *iScience*, 23(2), 100843. https://doi.org/10.1016/j.isci.2020.100843
- Kim, A. J., Fitzgerald, J. K., & Maimon, G. (2015). Cellular evidence for efference copy in drosophila visuomotor processing. *Nature Neuroscience*, 18(9), 1247–1255. https://doi.org/10.1038/ nn.4083
- Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *Proceedings of the National Academy of Sciences*, 111(24), 8997– 9002. https://doi.org/10.1073/pnas.1319438111
- Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory*, 14(9), 625–633. https://doi. org/10.1101/lm.663507
- Klaffehn, A. L., Baess, P., Kunde, W., & Pfister, R. (2019). Sensory attenuation prevails when controlling for temporal predictability of self- and externally generated tones. *Neuropsychologia*, 132, 107145. https://doi.org/10.1016/j.neuropsychologia.2019.107145
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1–16.
- Knapen, T., de Gee, J. W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., & Theeuwes, J. (2016). Cognitive and ocular factors jointly determine pupil responses under equiluminance. *PLoS One*, *11*(5), e0155574. https://doi.org/10.1371/journal.pone.0155574
- Knolle, F., Schröger, E., & Kotz, S. A. (2013a). Prediction errors in selfand 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
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3), 557–577. https://doi. org/10.1017/S0048577201990559
- Korka, B., Schröger, E., & Widmann, A. (2019). Action intentionbased and stimulus regularity-based predictions: Same or different? *Journal of Cognitive Neuroscience*, 31(12), 1917–1932. https://doi.org/10.1162/jocn\_a\_01456
- Korka, B., Schröger, E., & Widmann, A. (2020). What exactly is missing here? The sensory processing of unpredictable omissions is modulated by the specificity of expected action-effects. *European Journal of Neuroscience*, 52(12), 4667–4683. https:// doi.org/10.1111/ejn.14899
- Korka, B., Widmann, A., Waszak, F., Darriba, Á., & Schröger, E. (2021). The auditory brain in action: Intention determines predictive processing in the auditory system—A review of current paradigms and findings. *Psychonomic Bulletin & Review*, 29, 321–342. https://doi.org/10.3758/s13423-021-01992-z
- Krawczyk, M. C., Fernández, R. S., Pedreira, M. E., & Boccia, M. M. (2017). Toward a better understanding on the role of prediction error on memory processes: From bench to clinic. *Neurobiology* of *Learning and Memory*, *142*, 13–20. https://doi.org/10.1016/j. nlm.2016.12.011

GHTSLINK

- 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
- Lee, C. R., & Margolis, D. J. (2016). Pupil dynamics reflect behavioral choice and learning in a Go/NoGo tactile decision-making task in mice. *Frontiers in Behavioral Neuroscience*, 10, 200. https:// doi.org/10.3389/fnbeh.2016.00200
- Lee, M. D., & Wagenmakers, E.-J. (2013). Bayesian cognitive modeling: A practical course. Cambridge University Press. https://doi. org/10.1017/CBO9781139087759
- Lubinus, C., Einhäuser, W., Schiller, F., Kircher, T., Straube, B., & van Kemenade, B. M. (2021). Action-based predictions affect visual perception, neural processing, and pupil size, regardless of temporal predictability. *bioRxiv Neuroscience*. https://doi. org/10.1101/2021.02.11.430717
- MacDonald, P. A., & MacLeod, C. M. (1998). The influence of attention at encoding on direct and indirect remembering. *Acta Psychologica*, 98(2–3), 291–310. https://doi.org/10.1016/S0001 -6918(97)00047-4
- Makeig, S., Müller, M. M., & Rockstroh, B. (1996). Effects of voluntary movements on early auditory brain responses. *Experimental Brain Research*, 110(3), 487–492. https://doi.org/10.1007/BF002 29149
- Mama, Y., & Icht, M. (2016). Auditioning the distinctiveness account: Expanding the production effect to the auditory modality reveals the superiority of writing over vocalising. *Memory*, 24(1), 98–113. https://doi.org/10.1080/09658211.2014.986135
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods, 164(1), 177–190. https://doi.org/10.1016/j.jneum eth.2007.03.024
- Martikainen, M. H., Kaneko, K., & Hari, R. (2004). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, 15(3), 299–302. https://doi.org/10.1093/cercor/ bhh131
- Marzecová, A., Schettino, A., Widmann, A., SanMiguel, I., Kotz, S. A., & Schröger, E. (2018). Attentional gain is modulated by probabilistic feature expectations in a spatial cueing task: ERP evidence. *Scientific Reports*, 8(1), 54. https://doi.org/10.1038/ s41598-017-18347-1
- McGinley, M. J., David, S. V., & McCormick, D. A. (2015). Cortical membrane potential signature of optimal states for sensory signal detection. *Neuron*, 87(1), 179–192. https://doi.org/10.1016/j. neuron.2015.05.038
- 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
- Mifsud, N. G., Beesley, T., Watson, T. L., Elijah, R. B., Sharp, T. S., & Whitford, T. J. (2018). Attenuation of visual evoked responses to hand and saccade-initiated flashes. *Cognition*, 179, 14–22. https://doi.org/10.1016/j.cognition.2018.06.005
- Mifsud, N. G., & Whitford, T. J. (2017). Sensory attenuation of selfinitiated sounds maps onto habitual associations between motor action and sound. *Neuropsychologia*, 103, 38–43. https:// doi.org/10.1016/j.neuropsychologia.2017.07.019
- Mondor, T. A., & Morin, S. R. (2004). Primacy, recency, and suffix effects in auditory short-term memory for pure tones: Evidence from a probe recognition paradigm. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie*

*Expérimentale*, 58(3), 206–219. https://doi.org/10.1037/ h0087445

- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, 20(8), 750–756. https://doi.org/10.1016/j.cub.2010.02.045
- Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, 35(8), 4140–4154. https://doi.org/10.1002/hbm.22466
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24(4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x
- Naber, M., Frassle, S., Rutishauser, U., & Einhauser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of Vision*, *13*(2), 11. https://doi.org/10.1167/13.2.11
- Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, 15(7), 1040– 1046. https://doi.org/10.1038/nn.3130
- Nelson, A., & Mooney, R. (2016). The basal forebrain and motor cortex provide convergent yet distinct movement-related inputs to the auditory cortex. *Neuron*, 90(3), 635–648. https://doi. org/10.1016/j.neuron.2016.03.031
- Norman, K. A. (2010). How hippocampus and cortex contribute to recognition memory: Revisiting the complementary learning systems model. *Hippocampus*, 20(11), 1217–1227. https://doi. org/10.1002/hipo.20855
- 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
- Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. Proceedings of the National Academy of Sciences of the United States of America, 97(20), 11120–11124. https://doi. org/10.1073/pnas.97.20.11120
- Onton, J., & Makeig, S. (2006). Information-based modeling of eventrelated brain dynamics. *Progress in Brain Research*, *159*, 99–120. https://doi.org/10.1016/s0079-6123(06)59007-7
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 1–9. https://doi. org/10.1155/2011/156869
- Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719. https://doi.org/10.1016/ S1388-2457(00)00527-7
- Ott, C. G. M., & Jäncke, L. (2013). Processing of self-initiated speech-sounds is different in musicians. *Frontiers in Human Neuroscience*, 7, 41. https://doi.org/10.3389/fnhum.2013.00041
- Ozubko, J. D., Gopie, N., & MacLeod, C. M. (2012). Production benefits both recollection and familiarity. *Memory & Cognition*, 40(3), 326–338. https://doi.org/10.3758/s13421-011-0165-1
- Pantev, C., Eulitz, C., Hampson, S., Ross, B., & Roberts, L. E. (1996). The auditory evoked "off" response: Sources and comparison with the "on" and the "sustained" responses. *Ear and Hearing*,

17(3), 255–265. https://doi.org/10.1097/00003446-19960 6000-00008

PSYCHOPHYSIOLOGY SPR

- Pine, A., Sadeh, N., Ben-Yakov, A., Dudai, Y., & Mendelsohn, A. (2018). Knowledge acquisition is governed by striatal prediction errors. *Nature Communications*, 9(1), 1673. https://doi. org/10.1038/s41467-018-03992-5
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. Clinical Neurophysiology, 118(10), 2128–2148. https://doi. org/10.1016/j.clinph.2007.04.019
- Press, C., & Cook, R. (2015). Beyond action-specific simulation: Domain-general motor contributions to perception. *Trends in Cognitive Sciences*, 19(4), 176–178. https://doi.org/10.1016/j. tics.2015.01.006
- Press, C., Kok, P., & Yon, D. (2020). The perceptual prediction paradox. *Trends in Cognitive Sciences*, 24(1), 13–24. https://doi. org/10.1016/j.tics.2019.11.003
- Preuschoff, K., t' Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, 5, 115. https://doi. org/10.3389/fnins.2011.00115
- Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, *84*(2), 355– 362. https://doi.org/10.1016/j.neuron.2014.09.033
- Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, 7(1), 13289. https://doi. org/10.1038/ncomms13289
- Requarth, T., & Sawtell, N. B. (2011). Neural mechanisms for filtering self-generated sensory signals in cerebellum-like circuits. *Current Opinion in Neurobiology*, 21(4), 602–608. https://doi. org/10.1016/j.conb.2011.05.031
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64– 99). Appleton- Century-Crofts.
- Reznik, D., Henkin, Y., Schadel, N., & Mukamel, R. (2014). Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nature Communications*, 5(1), 4059. https://doi.org/10.1038/ncomm s5059
- Richer, F., & Beatty, J. (1987). Contrasting effects of response uncertainty on the task-evoked pupillary response and reaction time. *Psychophysiology*, 24(3), 258–262. https://doi.org/10.1111/ j.1469-8986.1987.tb00291.x
- Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences*, 24(2), 113–121. https://doi.org/10.1016/S0166 -2236(00)01685-4
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237. https://doi.org/10.3758/PBR.16.2.225
- Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. *Neuropsychologia*, 51(5), 922– 929. https://doi.org/10.1016/j.neuropsychologia.2013.02.005
- Roussel, C., Hughes, G., & Waszak, F. (2014). Action prediction modulates both neurophysiological and psychophysical indices of

sensory attenuation. *Frontiers in Human Neuroscience*, *8*, 115. https://doi.org/10.3389/fnhum.2014.00115

- Roy, J. E., & Cullen, K. E. (2001). Selective processing of vestibular reafference during self-generated head motion. *The Journal* of *Neuroscience*, 21(6), 2131–2142. https://doi.org/10.1523/ JNEUROSCI.21-06-02131.2001
- Rummell, B. P., Klee, J. L., & Sigurdsson, T. (2016). Attenuation of responses to self-generated sounds in auditory cortical neurons. *The Journal of Neuroscience*, 36(47), 12010–12026. https://doi. org/10.1523/JNEUROSCI.1564-16.2016
- SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology*, 50(4), 334–343. https://doi.org/10.1111/psyp.12024
- Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial suppression of self-generated sounds and its dependence on attention. *International Journal of Psychophysiology*, 90(3), 300–310. https://doi.org/10.1016/j.ijpsycho.2013.09.006
- Schafer, E. W. P., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science*, 181(4095), 175–177. https://doi.org/10.1126/science.181.4095.175
- Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160049. https://doi.org/10.1098/rstb.2016.0049
- Schiffer, A.-M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (2012). Surprised at all the entropy: Hippocampal, caudate and midbrain contributions to learning from prediction errors. *PLoS One*, 7(5), e36445. https://doi.org/10.1371/journ al.pone.0036445
- Schneider, D. M. (2020). Reflections of action in sensory cortex. Current Opinion in Neurobiology, 64, 53–59. https://doi. org/10.1016/j.conb.2020.02.004
- Schneider, D. M., & Mooney, R. (2018). How movement modulates hearing. Annual Review of Neuroscience, 41(1), 553–572. https:// doi.org/10.1146/annurev-neuro-072116-031215
- Schneider, D. M., Nelson, A., & Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature*, 513(7517), 189–194. https://doi.org/10.1038/nature13724
- Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. *European Journal of Neuroscience*, 41(5), 641–664. https://doi.org/10.1111/ejn.12816
- Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport*, 16(16), 1781–1785. https://doi.org/10.1097/01.wnr.0000185017.29316.63
- Simpson, H. M. (1969). Effects of a task-relevant response on pupil size. *Psychophysiology*, 6(2), 115–121. https://doi.org/10.1111/ j.1469-8986.1969.tb02890.x
- Sinclair, A. H., & Barense, M. D. (2019). Prediction error and memory reactivation: How incomplete reminders drive reconsolidation. *Trends in Neurosciences*, 42(10), 727–739. https://doi. org/10.1016/j.tins.2019.08.007
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative* and Physiological Psychology, 43(6), 482–489. https://doi. org/10.1037/h0055479

- Tapia, M. C., Cohen, L. G., & Starr, A. (1987). Attenuation of auditory-evoked potentials during voluntary movement in man. *International Journal of Audiology*, 26(6), 369–373. https://doi. org/10.3109/00206098709081565
- Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neuroscience*, 14(1), 2. https://doi. org/10.1186/1471-2202-14-2
- Timm, J., Schönwiesner, M., Schröger, E., & SanMiguel, I. (2016). Sensory suppression of brain responses to selfgenerated sounds is observed with and without the perception of agency. *Cortex*, 80, 5–20. https://doi.org/10.1016/j. cortex.2016.03.018
- Tonnquist-Uhlen, I., Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2003). Maturation of human central auditory system activity: The T-complex. *Clinical Neurophysiology*, *114*(4), 685– 701. https://doi.org/10.1016/S1388-2457(03)00005-1
- Twomey, D. M., Murphy, P. R., Kelly, S. P., & O'Connell, R. G. (2015). The classic P300 encodes a build-to-threshold decision variable. *European Journal of Neuroscience*, 42(1), 1636–1643. https:// doi.org/10.1111/ejn.12936
- Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications*, 8(1), 14637. https://doi.org/10.1038/ ncomms14637
- Van Slooten, J. C., Jahfari, S., Knapen, T., & Theeuwes, J. (2019). Correction: How pupil responses track value-based decision-making during and after reinforcement learning. *PLoS Computational Biology*, 15(5), e1007031. https://doi. org/10.1371/journal.pcbi.1007031
- Vinck, M., Batista-Brito, R., Knoblich, U., & Cardin, J. A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron*, *86*(3), 740–754. https://doi.org/10.1016/j.neuron.2015.03.028
- von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip: Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften*, 37(20), 464–476. https://doi. org/10.1007/BF00622503
- Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. Proceedings of the National Academy of Sciences of the United States of America, 97(20), 11125–11129. https://doi. org/10.1073/pnas.97.20.11125
- Williams, S. R., & Chapman, C. E. (2000). Time course and magnitude of movement-related gating of tactile detection in humans.
  II. Effects of stimulus intensity. *Journal of Neurophysiology*, 84(2), 863–875. https://doi.org/10.1152/jn.2000.84.2.863
- Williams, S. R., & Chapman, C. E. (2002). Time course and magnitude of movement-related gating of tactile detection in humans. III. Effect of motor tasks. *Journal of Neurophysiology*, 88(4), 1968–1979. https://doi.org/10.1152/jn.2002.88.4.1968
- Williams, S. R., Shenasa, J., & Chapman, C. E. (1998). Time course and magnitude of movement-related gating of tactile detection in humans. I. Importance of stimulus location. *Journal* of Neurophysiology, 79(2), 947–963. https://doi.org/10.1152/ jn.1998.79.2.947
- Wolpaw, J. R., & Penry, J. K. (1975). A temporal component of the auditory evoked response. *Electroencephalography* and Clinical Neurophysiology, 39(6), 609–620. https://doi. org/10.1016/0013-4694(75)90073-5

RIGHTSLINK4

- Wolpert, D., Ghahramani, Z., & Jordan, M. (1995). An internal model for sensorimotor integration. *Science*, 269(5232), 1880– 1882. https://doi.org/10.1126/science.7569931
- Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729–R732. https://doi.org/10.1016/S0960 -9822(01)00432-8
- Yebra, M., Galarza-Vallejo, A., Soto-Leon, V., Gonzalez-Rosa, J. J., de Berker, A. O., Bestmann, S., Oliviero, A., Kroes, M. C. W., & Strange, B. A. (2019). Action boosts episodic memory encoding in humans via engagement of a noradrenergic system. *Nature Communications*, 10(1), 3534. https://doi.org/10.1038/s41467-019-11358-8
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459–482. https://doi. org/10.1002/cne.920180503

#### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1 Supporting Information

**Figure S1.** Non-parametric cluster-based permutation test comparing the average EEG signal in the auditoryonly and the corrected motor-auditory condition (A– [MA–M]). Topographical maps denote the positive (red) and negative (blue) effects. The topography is shown for segments of 25 ms. The black dots indicate the electrodes over which the difference between the two conditions reaches significance. There were two significant clusters, a negative cluster (p < .001; 56–344 ms post-stimulus) and one positive cluster (p = .01; 122–232 ms post-stimulus)

How to cite this article: Paraskevoudi, N., & SanMiguel, I. (2023). Sensory suppression and increased neuromodulation during actions disrupt memory encoding of unpredictable self-initiated stimuli. *Psychophysiology*, *60*, e14156. <u>https://doi. org/10.1111/psyp.14156</u>