1 **Sensory suppression and increased neuromodulation during**

2 **actions disrupt memory encoding of unpredictable self-**

3 **initiated stimuli**

4

5 **Nadia Paraskevoudi[a,b] & Iria SanMiguel[a,b,c]**

6

7 *Affiliations:*

8 [a] Institute of Neurosciences, University of Barcelona, Barcelona, Spain

9 [b] Brainlab-Cognitive Neuroscience Research Group, Department of Clinical Psychology

10 and Psychobiology, University of Barcelona, Barcelona, Spain

11 [c] Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain.

12

13 *Corresponding author:*

14 Iria SanMiguel

15 Department of Clinical Psychology and Psychobiology

16 P. Vall d'Hebron 171

17 08035 Barcelona, Spain

18 Tel: (0034) 93 312 5050

19 Email: isanmiguel@ub.edu

20

21 *Declarations of interest:* **none.**

Abstract

Actions modulate sensory processing by attenuating responses to self- compared to externally-generated inputs, which is traditionally attributed to stimulus-specific 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 self-generation 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 – and 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.

Keywords: self-generation, memory, pupillometry, EEG, auditory processing

1. Introduction

44

45   Forming predictions about upcoming events in the environment is crucial for all behaving

46 organisms. A critical instance of such predictive processing is our ability to anticipate the

47 sensory consequences of our own actions, which is essential for building a sense of self and

48 shapes our perception of sense of agency (Gallagher, 2000). Although predictions have been

49 suggested to facilitate perceptual processing in the wider sensory literature, most studies from

50 the action domain point to attenuated processing for the predicted self-produced events

51 (Schröger et al., 2015; Press et al. 2020), with only a few exceptions showing the opposite

52 effect (e.g., Reznik et al., 2014; Eliades & Wang, 2008). Several lines of research have shown

53 that actions suppress the processing of the self-generated reafferent input (so-called self-

54 generation effect) in a wide range of species (Chagnaud et al., 2015; Kelley et al., 2010; Kim

55 et al., 2015; Requarth & Sawtell, 2011; Roy et al., 2001; Schneider et al., 2014) and

56 irrespective of sensory modality (auditory; Baess et al., 2011; Horváth, 2013a; Horváth,

57 2013b; Martikainen, 2004; Mifsud & Whitford, 2017; SanMiguel et al., 2013; Saupe et al.,

58 2013; Schafer & Marcus, 1973; Timm et al., 2013; Klaffehn et al., 2019; Weller et al., 2016;

59 Pyasik et al., 2018, visual; Hughes & Waszak, 2011; Mifsud et al., 2018; Roussel et al., 2013;

60 Roussell et al., 2014, and tactile; Blakemore et al., 1998; Hesse et al., 2010; Kilteni et al.,

61 2020). Dominant *cancellation* models attribute this attenuation to stimulus-specific

62 prediction signals generated via internal forward modelling (Wolpert et al., 1995) before or

63 during an action that are sent from the motor to the corresponding sensory cortices (Sperry,

64 1950; von Holst, 1954). These motor-induced predictions of sensory reafference (i.e.,

65 corollary discharge) are compared to the sensory input generated by one's actions, and only

66 the difference between the two (i.e., prediction error) is sent to higher stages of the neuronal

67 hierarchy for further processing (Friston, 2005), ultimately suppressing the processing of the

68 anticipated event in order to prioritize the most informative unexpected inputs (Sperry, 1950;

69 von Holst, 1954).

70   So far, self-generation effects in the auditory domain have been typically assessed using

71 the *contingent* self-generation paradigm, where neural responses to sounds generated by the

72 participants in a fully predictable fashion are compared to the responses elicited by

73 externally-generated sounds. Most of these studies have shown attenuated auditory N1 and

74 P2 event-related potential (ERP) amplitudes, with especially the former being considered a

75 proxy of suppressed processing of self-produced sounds in the auditory cortex, driven by
76 motor predictions (for a review see Schröger et al., 2015). However, it is known that the
77 auditory N1 response is not a unitary phenomenon but rather reflects the overlap of several
78 components (SanMiguel et al., 2013), among which two are proposed to be stimulus-specific
79 and to reflect processing in primary and secondary auditory areas (Näätänen & Picton, 1987).
80 The first one is generated by tangentially oriented sources in the auditory cortex, has a
81 frontocentral distribution, and shows polarity reversal at the mastoids. Nevertheless, the few
82 studies that have analyzed N1 amplitudes at the mastoids have reported no suppression
83 (Timm et al., 2013; SanMiguel et al., 2013) or even enhanced amplitude in response to self-
84 generated sounds (Horváth et al., 2012). The second one, usually referred to as the "T
85 complex", is generated by radial sources in the superior temporal gyrus and is typically
86 identified as the first and second negative peaks (i.e., Na and Tb, respectively) observable on
87 anterior temporal sites (Tonnquist-Uhlen et al., 2003; Wolpaw & Penry, 1975). Reports of
88 self-generation effects on the "T complex" remain scarce, but the few studies that have
89 assessed it reported attenuation of Tb for self-generated sounds (Horváth, 2013b; SanMiguel
90 et al., 2013). Inevitably, if N1-suppression indeed reflects modulations in auditory areas
91 driven by stimulus-specific motor predictions, then the self-generation effects should be
92 specific to the expected stimulus and mediated by sensory specific areas.

93 However, the stimulus-specificity of the effects has been challenged by data showing that
94 the N1-suppression may mostly affect the stimulus-unspecific component of N1 (SanMiguel
95 et al., 2013), which is thought to be the cortical projection of a reticular process that facilitates
96 motor activity (Näätänen & Picton, 1987). Further evidence supporting the lack of stimulus-
97 specificity of the effects comes from work showing that suppression of responses can be also
98 observed in the absence of a predictive relationship between the action and the sound.
99 Horváth et al. (2012) employed a *coincidence* rather than a *contingent* task, where
100 participants had to press a button several times and concurrently, but independently from the
101 actions, a sound sequence with random between-sound intervals was presented. They showed
102 that despite the absence of contingent relationship between the action and the sound, the
103 auditory N1 was attenuated for sounds that coincided with a button press, indicating that the
104 N1-suppression for self-produced sounds may be also driven by the temporal proximity
105 between action and sound, rather than by a stimulus-specific prediction of the expected

106   sensory input. These findings challenge the predictions made by the cancellation account,

107   since if this attenuation was to be attributed to stimulus-specific motor predictions, then it

108   should only be found in paradigms where the stimulus can be indeed predicted by the action,

109   and it should mainly reflect the attenuation of the sensory components of N1 (generated in

110   auditory cortex).

111   Indeed, a significant number of non-predictive processes are also known to modulate

112   perceptual and neural responses during motor acts (Press et al., 2020; Press & Cook, 2015),

113   which may point to an unspecific gating mechanism during movement (i.e., a "halo of

114   neuromodulation") that modulates processing of sounds presented close in time to a motor

115   act even in the absence of a causal relationship between the action and the sensory stimulus

116   (e.g., Hazemann et al., 1975; Makeig et al., 1996; Horváth, 2013a, b; Horváth et al., 2012).

117   This mechanism may be mediated by arousal-related unspecific modulatory processes which

118   receive influences from motor areas (Näätänen & Picton, 1987). Specifically, the locus

119   coeruleus norepinephrine (LC-NE) system has been implicated as a possible reticular

120   candidate responsible for mediating the unspecific gating of sensory processing around the

121   motor acts. Supporting evidence comes from animal work showing that the auditory cortex

122   receives inputs from both motor and neuromodulatory areas (mostly from the basal

123   forebrain), which are simultaneously active during movement and form synapses onto many

124   of the same auditory cortical excitatory and inhibitory neurons (Nelson & Mooney, 2016; for

125   a review see Schneider & Mooney, 2018). Critically, the neurons in the basal forebrain

126   receive inputs from subcortical regions, including the locus coeruleus, suggesting that this

127   overlap of motor and neuromodulatory inputs in auditory areas may result in a diverse set of

128   motor and neuromodulatory influences, thereby pointing to coexisting, but possibly

129   independent, stimulus-specific and unspecific effects during movement (Nelson & Mooney,

130   2016). More importantly, the link between LC-NE activity and motor behavior is supported

131   by animal (Lee & Margolis, 2016; McGinley et al., 2015; Stringer et al., 2019; Vinck et al.,

132   2015) and human work (Simpson, 1969; Strauch et al., 2020; Yerba et al., 2019) showing a

133   close association between engaging in motor activities (e.g., whisking or button press in

134   Go/No-Go tasks) and pupil dilation, which has been shown to track the activity of the LC-

135   NE system (Aston-Jones & Cohen, 2005; Vinck et al., 2015).

136   Although most studies have focused on the effects of self-generation on the immediate
137   sensory processing, previous work has reported modulatory effects of movements on
138   hippocampal and parahippocampal activity (Halgren, 1991; Mukamel et al., 2010; Rummell
139   et al., 2016), suggesting that the differential processing of self- and externally-generated
140   stimuli may also have consequences for memory encoding. One line of evidence shows that
141   spoken words are better remembered than words that are passively listened to (MacDonald
142   & MacLeod, 1998) and played melodies can be better recognized compared to melodies that
143   are passively presented to the participants (Brown & Palmer, 2012), suggesting that self-
144   generated stimuli are encoded more efficiently in memory than the passively presented ones
145   (so-called *production effect*). These memory improvements for self-produced stimuli have
146   been attributed to the increased distinctiveness of those items because producing them
147   provides extra mnemonic information (i.e., a memory trace of having spoken the words) that
148   is not present for silently read words (Conway & Gathercole, 1987; Mama & Icht, 2016;
149   Ozubko et al., 2012). Additional evidence suggests that the action-induced memory
150   enhancement may be driven by the engagement of the noradrenergic system, as shown by
151   the increased pupil dilation and locus coeruleus activity in response to stimuli tied with – but
152   not produced by – actions (i.e., Go-events in a Go/No-Go task; Yerba et al., 2019). However,
153   predictive coding theories suggest that learning and memory are driven by the amount of
154   surprise (i.e., prediction error) associated with an item (Bar, 2009; Krawczyk et al., 2017;
155   Pine et al., 2018). Specifically, the larger prediction errors elicited by unpredicted items at
156   encoding are thought to result in greater synaptic change, reduced prediction error for this
157   item at retrieval (Greve et al., 2017; Heilbron & Chait, 2018; Henson & Gagnepain, 2010;
158   Pine et al., 2018), and ultimately enhanced memory performance. The latter findings are
159   consistent with studies reporting that items leading to high prediction error tend to produce
160   greater hippocampus fMRI signal in the study phase of a recognition memory experiment,
161   and that this increase in hippocampal activity is associated with enhanced subsequent
162   recollection in the test phase (Gagnepain et al., 2011; Henson & Gagnepain, 2010; Pine et
163   al., 2018). In sum, this framework would predict memory enhancements for the externally-
164   generated sounds in a typical *contingent* paradigm where they inherently elicit larger
165   prediction error compared to the more predictable self-generated stimuli.

166   To the best of our knowledge, there have been no attempts to simultaneously address the

167   effects of self-generation on sensory processing and memory encoding of sounds and assess

168   the possible relationship between these two phenomena. Meanwhile, the findings as to the

169   direction of the motor-driven effects on memory encoding of self-generated stimuli

170   (enhanced vs. reduced memory performance) remain mixed but point to differential memory

171   representations for events that have been experienced as self-initiated and those that have

172   been experienced as externally-generated. Further, recent evidence contradicts the traditional

173   view of sensory suppression being due to stimulus-specific motor predictions, thereby raising

174   the need to examine the possible role of stimulus-unspecific neuromodulatory mechanisms

175   both on sensory processing and memory encoding of sounds merely coinciding with motor

176   actions.

177   In this study, we examined whether and how motor actions affect sensory processing and

178   memory encoding of concomitant, but unpredictable sounds, by employing a combination of

179   a self-generation and memory recognition task, while monitoring the brain's and the pupil's

180   responses to sounds that were either presented passively or that coincided in time with a

181   motor act. The aim of this study was twofold: First, we aimed to clarify the role of the

182   neuromodulatory LC-NE system in the motor-driven modulation of auditory processing of

183   self-generated sounds. Based on previous work, we expected to replicate the sensory

184   suppression effects when sounds merely coincide with, rather than being predicted by the

185   action. Thus, we expected to observe typical self-generation effects at encoding (i.e.,

186   attenuated N1, P2, Tb responses and enhanced P3, Hórvath et al., 2012; SanMiguel et al.,

187   2013). Moreover, we hypothesized that these typical self-generation effects are related to

188   neuromodulatory processes, thus we expected them to correlate with increased pupil diameter

189   for motor-auditory events (McGinley et al., 2015; Yerba et al., 2019). Second, we sought out

190   to examine whether the differential sensory processing of stimuli paired with an action affects

191   their encoding in memory. To this end, we examined whether coincidence with an action at

192   encoding enhances or impairs memory recall at retrieval, and we hypothesized that the

193   potential differences in the memory encoding of sounds presented with or without a

194   concomitant action should by driven by, and thus correlate with, the differential

195   neurophysiological responses (i.e., event-related potentials and pupil diameter) at encoding

196   for sounds that were either paired with an action or not.

2. Methods

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 twenty-three 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 the 48 hours 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.

Encoding phase: At the start of each trial, subjects were presented with a row of six vertical lines on the screen, separated in semi-random 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 (12 seconds), 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 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 event; MA). The other half did not result in the presentation of a sound (Motor-only 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 nine events, three motor-only (M) and six sounds (i.e., three of them triggered by a button and the other three presented passively between presses; MA and A events, respectively). The event-to-event onset asynchrony varied from 0.8 s to 2.4 s, in steps of 0.05

228   s, while the sound-to-sound onset asynchrony varied between 1.6 and 2.4 seconds in steps of

229   0.05. Participants were asked to pay attention to all the sounds presented. The encoding phase

230   finished when the horizontal line had intersected all the vertical ones and arrived at the right

231   of the screen. If the task was performed correctly (i.e., all required button presses were

232   performed), the trial continued into the retention phase. Otherwise, an error message

233   appeared on the screen indicating that the participant did not press the button every time the

234   horizontal line reached a vertical one, and a new trial began.

235   *Retention phase:* During the subsequent retention phase, participants were presented with

236   a fixation cross on the screen for 3 s and they were asked to remember all the sounds that

237   were previously presented in the encoding phase.

238   *Retrieval phase:* In the retrieval phase, participants were presented with two sounds with

239   a 2 s sound-to-sound onset asynchrony (indicated by the visual stimuli "Sound 1" and "Sound

240   2", respectively). Subsequently, a question appeared on the screen, prompting participants to

241   respond whether the first or the second test sound was presented during the encoding phase.

242   The response window was 2 seconds. After the end of the response window, a fixation was

243   presented for 2 seconds (inter-trial interval) before the start of the next trial.

244

245   **2.3. Stimuli**

246   **2.3.1. Sequences**

247   Two types of sequences were created, differing in whether both or only one of the test

248   sounds presented at retrieval were also present during the encoding phase. In the "Two Test

249   Sounds at Encoding" (henceforth 2T; Figure 1) and unbeknownst to the participants, the two

250   test sounds presented passively at retrieval were also presented in the encoding sequence,

251   one as a motor-auditory (Encoded as MA) and the other one as an auditory-only event

252   (Encoded as A). In the "One Test Sound at Encoding" (henceforth, 1T; Figure 1), only one

253   of the test sounds at retrieval was presented at encoding, either as a MA (Encoded as MA) or

254   as an A event (Encoded as A), while the other sound was new (New sound). The 1T

255   sequences were introduced only for the behavioral data and were not used for the EEG and

256   pupillometry analyses. This design allowed us to have enough trials for Encoded as A and

257   Encoded as MA sounds at retrieval, keep the experiment's duration within a reasonable time,

258   and obtain an additional objective measure of memory performance in the 1T sequences

259   besides the measure of memory bias obtained in the 2T sequences. Five of the 1T sequences

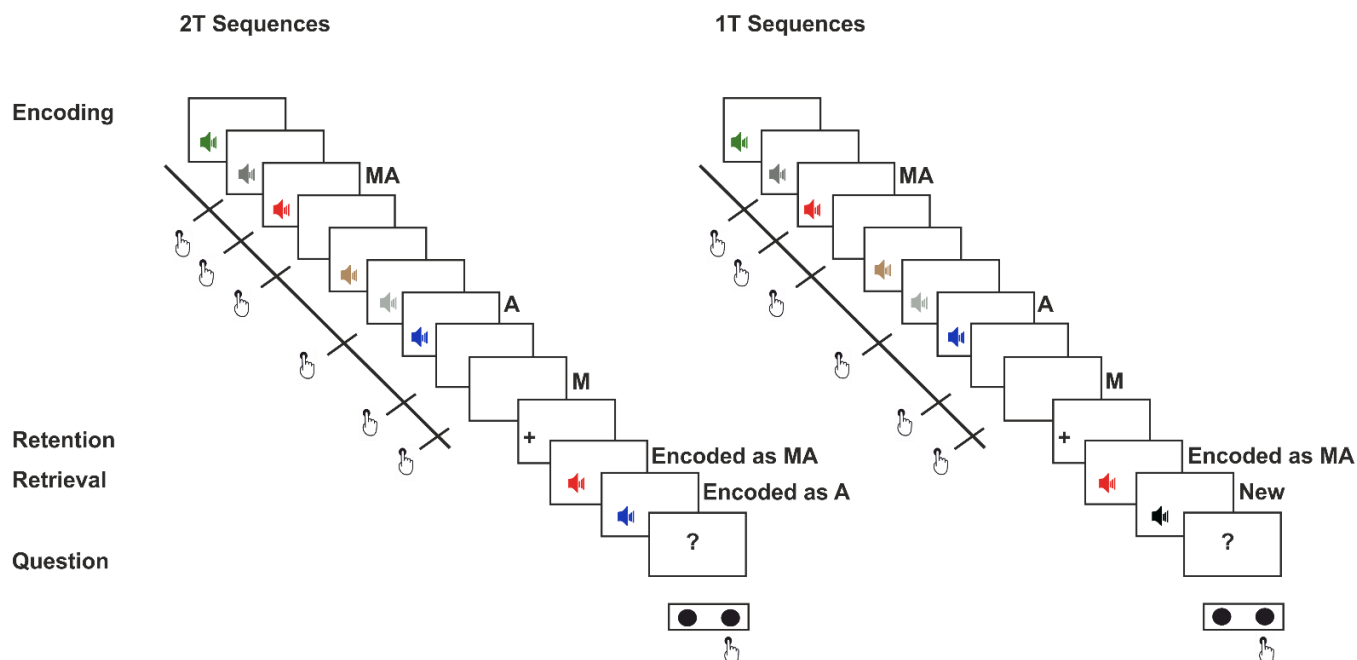260   were randomly chosen to be used during the practice block.

261

262

263   **Figure 1**. Schematic representation of the experimental design. Each trial consisted of three
264   phases (encoding, retention, retrieval). During the encoding phase, a horizontal line moved
265   at a stable pace across the screen from left to right, intersecting each of the six vertical lines
266   as it advanced. Participants were asked to press a button every time the horizontal line
267   reached one of the vertical ones. Only half of these presses produced a sound (Motor-
268   auditory; MA). The other half did not result in the presentation of a sound (Motor-only; M).
269   Additionally, three more sounds were presented passively to the participants without being
270   triggered by a button press (Auditory-only condition; A). The encoding phase finished when
271   the horizontal line had intersected all the vertical ones and arrived at the right of the screen.
272   During the subsequent retention phase, participants were presented with a fixation cross on
273   the screen for 3 s and they were asked to remember all the sounds that were previously
274   presented in the encoding phase. In the retrieval phase, we employed two types of sequences:
275   2T (left) and 1T (right) sequences. Participants were presented with two sounds separated by
276   a 2 s sound-to-sound onset asynchrony. In the 2T sequences, both sounds at retrieval were
277   presented also at encoding, one Encoded as A and the other Encoded as MA. In the 1T
278   sequences, only one of the two sounds was presented before, either Encoded as A or Encoded
279   as MA, while the other sound was new. After the presentation of the test sounds, a question
280   appeared on the screen, prompting participants to respond whether the first or the second test
281   sound was presented during the encoding phase. The response window was 2 s and the inter-
282   trial interval was set to 2 s.

283

284   Importantly, the same sounds in the same encoding sequence positions were used as

285   either A or MA in different trials, which allowed us to compare between physically identical

286   auditory sequences that only differed in the actions performed. Additionally, we controlled

287   for the order of the sounds at encoding that would be later used as test at retrieval for the 2T

288   sequences, the order of the two retrieval sounds, and the position of the test sounds in the

289   encoding sequence counterbalancing it across trials. Specifically, related to the latter, we

290   discarded the first and last position of the encoding sequence for placing test sounds to avoid

291   primacy and recency effects, which refer to an improvement in memory retention for stimuli

292   that have been presented first or last in a list, respectively (Mondor & Morin, 2004).

293   However, we included 20 catch trials with test sounds in those positions, which were

294   randomly interleaved with the experimental sequences described above and discarded from

295   all analyses.

296

297   **2.3.2. Auditory stimuli**

298   For the main experiment, 255 different, environmental, natural, complex, and non-

299   identifiable sounds were gathered from the libraries of McDermott Sound Library

300   (http://mcdermottlab.mit.edu/svnh/Natural-Sound/Stimuli.html) and Adobe

301   (https://offers.adobe.com/en/na/audition/offers/audition_dlc.html). These sounds were then

302   edited to all have 250 ms of duration, a sampling rate of 44.1 kHz and to be played at 16 bits,

303   mono and with 50-55 dB of intensity. Subsequently, six volunteers that did not participate in

304   the main experiment rated the 255 sounds based on their identifiability (i.e., how easy it was

305   to assign a name to them). All sounds were presented to them in a randomized order and each

306   sound was presented twice. The volunteers rated them in a scale from 1-3 (1 = identifiable,

307   2 = not sure, 3 = not identifiable), and the mean score for each sound was calculated. The

308   108 less identifiable sounds were selected to construct the unique experimental sound

309   sequences. The sounds used in the practice block consisted of 35 pure tones of different

310   frequencies, ranging from 300 Hz to 3700 Hz in steps of 100.

311

312   **2.4. Apparatus**

313   The visual stimuli were presented on an ATI Radeon HD 2400 monitor. The auditory

314   stimuli were presented via the Sennheiser KD 380 PRO noise cancelling headphones. To

315 record participants' button presses and behavioural responses, we used the Korg nanoPAD2.

316 The buttons of this device do not produce any mechanical noise when pressed, and, thus, do

317 not interfere with our auditory stimuli. The presentation of the stimuli and recording of

318 participants' button presses and responses were controlled using MATLAB R2017a (The

319 Mathworks Inc., 2017), the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al.,

320 2007; Pelli, 1997), and the Eyelink add-in toolbox for eyetracker control.

321 EEG activity was acquired with Neuroscan 4.4 software and Neuroscan SynAmps RT

322 amplifier (NeuroScan, Compumedics, Charlotte, NC, USA). We recorded continuously with

323 Ag/AgCl electrodes from 64 standard locations according to the 10% extension of the

324 International 10–20 system (Chatrian, Lettich, & Nelson, 1985; Oostenveld & Praamstra,

325 2001) mounted in a nylon cap (Quick-Cap; Compumedics, Charlotte, NC, USA). An

326 additional electrode was placed at the tip of the nose (serving as online reference). The

327 vertical electrooculogram (EOG) was measured with two electrodes placed above and below

328 the left eye, and the horizontal EOG with two electrodes placed on the outer canthi of the

329 eyes referenced to the common reference (unipolar montage). The ground electrode was

330 placed at AFz. All impedances were kept below 10 kΩ during the whole recording session

331 and data was sampled at 500 Hz.

332 Concurrently with the EEG recording, horizontal and vertical gaze position, as well as the

333 area of the pupil, were recorded using EyeLink 1000 desktop mount (SR Research, sampling

334 rate: 1,000 Hz; left eye recordings except for three participants for whom the right eye was

335 recorded instead). The pupil was assessed in the centroid mode of the eye tracker, which uses

336 a center-of-mass algorithm. This algorithm detects the pupil area by identifying the number

337 of black pixels and its center on the video image. Importantly, in contrast to methods using

338 ellipse fitting for the measurement of the pupil, this method is hardly affected by noise (S-R

339 Research Eyelink-CL Manual, p. 71).

340

341 **2.5. Procedure**

342 Prior to the start of the experiment, participants were asked to complete several

343 questionnaires. Subsequently, participants were seated in an electrically and acoustically

344 shielded room and were asked to place their head on a chinrest at approximately 60 cm from

345 the screen. Eyetracker calibration was performed first at the start of the experiment and then

346 every six blocks. In order to familiarize themselves with the task, participants first completed

347 a practice block of 5 trials and repeated it as many times as they needed to make sure they

348 understood how to perform the task. During the main experiment, participants completed a

349 total of 236 trials: 56 1T trials, 160 2T trials and 20 catch trials. These were divided in 24

350 blocks, 20 of them comprised of 10 trials (9 experimental and 1 catch trial) and the remaining

351 4 comprised of 9 trials (all of them experimental trials). At the end of each block, a message

352 appeared informing participants about the number of errors (i.e., not pressing the button when

353 required) and extra-presses (i.e., more than the required button presses) at the encoding phase,

354 as well as the number of missed responses at retrieval for this block. Catch trials, as well as

355 trials including errors in button-pressing and missed responses were discarded from further

356 analyses. Participants took a break of approximately 5 minutes every six blocks to prevent

357 fatigue. The experiment lasted for approximately 1.5 hour excluding the EEG preparation.

358

359 **2.6. Data analysis**

360 **2.6.1. Behavioral analysis**

361 To test for differences in memory bias (2T sequences) and memory performance (1T

362 sequences) for sounds encoded as A or MA, we performed two different analyses. For the 1T

363 sequences, we calculated the percent correct for the sounds at retrieval (i.e., memory

364 performance), separately for those that were Encoded as MA and Encoded as A, which was

365 subsequently submitted to a two-sided paired sample t-test. For the 2T-trial sequences, we

366 calculated the percent recall for sounds Encoded as MA and Encoded as A and tested for

367 differences in memory bias, using a two-sided paired samples t-test. We complemented the

368 frequentist t-tests with corresponding Bayesian t-tests, separately for the 1T and 2T

369 sequences. For both Bayesian comparisons, the Bayes factor (BF_{10}) for the alternative

370 hypothesis (i.e., the difference of the means is not equal to zero) was calculated. Specifically,

371 the null hypothesis corresponded to a standardized effect size $\delta = 0$, and the alternative

372 hypothesis was defined as a Cauchy prior distribution centered around 0 with a scaling factor

373 of $r = 0.707$ (Rouder et al., 2012). In line with the Bayes Factor interpretation (Lee &

374 Wagenmakers, 2013) and with previous work reporting Bayes Factors (Korka et al., 2019;

375 Korka et al., 2020; Marzecová et al., 2018), data were taken as moderate evidence for the

376 alternative hypothesis if the BF_{10} was greater than 3, while values close to 1 were considered

377 only weakly informative. Values greater than 10 were considered strong evidence for the

378 alternative (or null) hypothesis.

379

380 **2.6.2. EEG preprocessing**

381 EEG data was analyzed with EEGLAB (Delorme & Makeig, 2004) and plotted with

382 EEProbe (ANT Neuro). Data were high-pass filtered (0.5 Hz high-pass, Kaiser window,

383 Kaiser β 5.653, filter order 1812), manually inspected so as to reject atypical artifacts and

384 identify malfunctioning electrodes, and corrected for eye movements with Independent

385 Component Analysis, using the compiled version of runica (binica) that uses the logistic

386 infomax ICA algorithm (Onton & Makeig, 2006). Components capturing eye movement

387 artifacts were rejected by visual inspection and the remaining components were then

388 projected back into electrode space. Data was then low-pass filtered (30 Hz low-pass, Kaiser

389 window, Kaiser β 5.653, filter order 1812), remaining artifacts were rejected by applying a

390 75 µV maximal signal-change per epoch threshold, and malfunctioning electrodes were

391 interpolated (spherical interpolation). A −100 to +500 ms epoch was defined around each

392 event both at the encoding and the retrieval phase. The data was subsequently baseline

393 corrected (100 ms prior to the event). We calculated the average wave for each event of

394 interest, as well as the grand average for the whole sample. Specifically, we obtained the

395 averages for the MA, A, and M events at encoding, while for the retrieval data, we binned

396 the responses to motor-auditory and auditory-only sounds as a function of memory (i.e.,

397 Encoded as MA and Encoded as A at retrieval that were remembered or forgotten). For each

398 condition of interest the number of remaining trials used for the analyses after trial rejection

399 were: Auditory-only ($M = 424.9$, $SD = 46.9$), Motor-auditory ($M = 427.2$, $SD = 40.6$), Motor-

400 only ($M = 429$, $SD = 40.8$), Encoded as A and forgotten ($M = 68$, $SD = 11.7$), Encoded as A

401 and remembered ($M = 64$, $SD = 14.7$), Encoded as MA and forgotten ($M = 64.1$, $SD = 14.2$),

402 Encoded as MA and remembered ($M = 67.7$, $SD = 11.9$).

403 To assess self-generation effects at encoding, MA sound responses were corrected for

404 motor activity by subtracting the motor-only (M) averages from the motor-auditory (MA)

405 averages, since the signal obtained in the MA condition represents the brain signal elicited

406 by the sound, but also by the planning and execution of the finger movement to press the

407 button. We, therefore, obtained a motor-corrected wave that only included the brain signal

408   elicited by the MA sound. Self-generation effects at encoding were then assessed by

409   comparing responses to MA sounds corrected for motor activity (MA–M) with the responses

410   elicited by the auditory-only sounds (A). Self-generation effects are presented in all figures

411   as the difference wave between the motor-auditory (corrected) sound responses and the

412   auditory-only sound responses (A–[MA–M]). No motor correction was performed at

413   retrieval since the Encoded as MA sounds were presented passively.

414

2.6.3. ERP analysis

416   For all the effects of interest at encoding, we examined responses separately for the N1

417   and P2 at Cz (N1, P2) and at the mastoids (henceforth, $N1_{mast}$, $P2_{mast}$), the P3 component at

418   Pz, and the N1 subcomponents Na and Tb at T7 and T8. The same components except for P3

419   were examined at retrieval. The windows were defined after visual inspection of the data by

420   locating the highest negative or positive (depending on the component of interest) peak in

421   the usual latencies for each component as reported by previous work (SanMiguel et al., 2013).

422   Specifically, time windows for N1 (and $N1_{mast}$), P2 (and $P2_{mast}$), Na, and Tb were defined on

423   the grand-averaged waveforms of the auditory-only sounds as previously reported (e.g.,

424   SanMiguel et al., 2013). Na and Tb were identified as the first and second negative peaks,

425   respectively, identifiable after sound onset on electrodes T7 and T8, as recommended by

426   Tonnquist-Uhlen et al. (2003). $N1/N1_{mast}$ and $P2/P2_{mast}$ were identified as the negative and

427   positive peaks occurring in the window ~70 to 150 ms, and ~150 to 250 ms after stimulus

428   onset on Cz, respectively, showing reversed polarity at the mastoid electrodes. P3 at encoding

429   was identified as the peak of the difference wave (A – [MA-M]) in the P3 window range

430   based on previous work (e.g., Baess et al., 2008). The time windows for the $N1/N1_{mast}$,

431    $P2/P2_{mast}$, P3, Na, and Tb peaks were centered on the identified peaks ± 13, 25, 15, 10, and

432   15 ms, respectively. For the encoding data, we performed paired samples *t*-tests with the

433   factor Sound Type (A vs. MA) to test for differences in N1, P2 and P3, and a repeated

434   measures ANOVA with factors Sound Type (A vs. MA) x Laterality (M1 vs. M2 or T7 vs.

435   T8) to test for differences in $N1_{mast}$, $P2_{mast}$ and Na, Tb, respectively. For the retrieval data we

436   performed 2x2 ANOVAs with the factors Sound Type (Encoded as A vs. Encoded as MA)

437   and Memory (Remembered vs. Forgotten) on N1 and P2, while for the $N1_{mast}$, $P2_{mast}$, Na, and

438   Tb an additional factor Laterality was introduced in the ANOVAs (i.e., M1 vs. M2 or T7 vs.

439   T8).

440

441   ### 2.6.4. Pupillometry analysis

442   Missing data and blinks, as detected by the EyeLink software, were padded by 100 ms and

443   linearly interpolated. Additional blinks were found using peak detection on the velocity of

444   the pupil signal and linearly interpolated (Urai et al., 2017). Blinks separated by less than 250

445   ms were aggregated to a single blink. The interpolated pupil time series were bandpass

446   filtered using a 0.05–4 Hz third-order Butterworth filter. Low-pass filtering reduces

447   measurement noise not likely to originate from physiological sources, as the pupil functions

448   as a low-pass filter on fast inputs (Binda et al., 2013; Hoeks & Levelt, 1993). High-pass

449   filtering removes slow drifts from the signal that are not accounted for by the model in the

450   subsequent deconvolution analysis. First, we estimated the effect of blinks and saccades on

451   the pupil response through deconvolution and removed these responses from the data using

452   linear regression using a procedure detailed in previous work (Knapen et al., 2010; Urai et

453   al., 2017). The residual bandpass filtered pupil time series was used for the evoked analyses

454   (van Slooten et al., 2018). After zscoring per trial, we epoched trials (epoching window -0.5

455   to 1.5 post-event), baseline corrected each trial by subtracting the mean pupil diameter

456   500 ms before onset of the event and resampled to 100 Hz.

457   For each participant, we first obtained the average evoked response for the main events of

458   interest. Specifically, we obtained the averages for the A and MA events at encoding, while

459   at retrieval we obtained the averages for the Encoded as A and Encoded as MA sounds,

460   separately for the remembered and the forgotten ones. We used non-parametric permutation

461   statistics to test for the group-level significance of the individual averages, separately for

462   encoding and retrieval. Specifically, we computed t values of the difference between the two

463   conditions of interest and thresholded these t values at a p value of 0.05. Each cluster was

464   constituted by the samples that passed the threshold of the p value. The cluster statistics was

465   chosen as the sum of the paired t-values of all the samples in the cluster. First, we compared

466   the pupil response to motor-auditory and auditory-only events at encoding. For the retrieval

467   data, we aimed to test for possible main effects of Sound Type (Encoded as A vs. Encoded

468   as MA) and Memory (Remembered vs. Forgotten), as well as for possible interactions

469 between the two. For the main effects of Sound Type and Memory at retrieval, the

470 permutation statistics were performed between Encoded as A and Encoded as MA sounds

471 (irrespective of their memory) and between Remembered and Forgotten sounds (irrespective

472 of how they were encoded before), respectively. To test for possible interactions, the cluster-

473 permutation test was performed on the difference waves ([Encoded as A and remembered –

474 Encoded as MA and remembered] and [Encoded as A and forgotten – Encoded as MA and

475 forgotten]). For each statistical test, this procedure was performed by randomly switching

476 labels of individual observations between these paired sets of values. We repeated this

477 procedure 10,000 times and computed the difference between the group means on each

478 permutation. The obtained p value was the fraction of permutations that exceeded the

479 observed difference between the means (i.e., two-sided dependent samples tests). The pupil

480 preprocessing and analysis was performed with custom software based on previous work

481 (Urai et al., 2017) using Fieldtrip (Oostenveld et al., 2011).

482

2.6.5. Correlations

484 Finally, we hypothesized that the electrophysiological and neuromodulatory effects at

485 encoding (i.e., sensory suppression and pupil dilation for MA events) might be driving any

486 memory encoding differences between A and MA sounds, and that neuromodulation might

487 be behind the suppression of ERP responses to MA sounds. To assess these relationships, we

488 tested for possible correlations between the behavioural, electrophysiological and

489 neuromodulatory (i.e., pupil diameter) effects of actions. Only those differences between MA

490 and A events that were found to be significant in the previous analyses were introduced in

491 the correlation analyses. For all the behavioural and the electrophysiological effects, we first

492 calculated the difference by subtracting the MA from A values (i.e., difference in memory

493 and ERP amplitude for each component of interest between A and MA). Regarding the ERPs

494 identified in two electrodes (e.g., Na, Tb, $N1_{mast}$, $P2_{mast}$), we calculated the mean amplitude

495 across the two (T7/T8 and M1/M2, respectively). For the pupil data, we used the peak of the

496 difference wave between A and MA events at encoding. We then submitted these values to

497 a Pearson correlation coefficient to test for correlations between a) the effects on ERPs at

498 encoding and memory performance/bias (1T and 2T sequences, respectively), b) the

499 neuromodulatory effects at encoding and memory performance/bias (1T and 2T sequences,

500 respectively), and c) the effects on the ERPs and the neuromodulatory effects at encoding. In

501 all correlations, for the ERPs, the larger the attenuation effects for the negative (N1, P2$_{mast}$,

502 Na, Tb) and positive (N1$_{mast}$, P2, P3) components, the more negative and positive the values,

503 respectively. Conversely, for the pupil and the behavioural data, the more negative the value,

504 the larger the pupil diameter and the worse the memory performance for MA sounds.

505

506 ### 3. Results

507 All statistical analyses were performed using R (version 3.6.0). For all the *t*-tests

508 performed, we first confirmed that the assumption of normality was not violated (Shapiro–

509 Wilk normality test $p > .05$).

510

511 **3.1. Behavioural performance**

512 For the analysis of the behavioural data, we calculated the percent correct (i.e., memory

513 performance in the 1T sequences) and the percent recall (memory bias in the 2T sequences)

514 for sounds that were encoded as motor-auditory or auditory-only (see Figure 2). For the 1T

515 sequences, we obtained significantly better memory performance for sounds that were

516 encoded as auditory-only compared to those that coincided with participants' motor acts in

517 the previous encoding phase, $t(22) = 3.15$, $p = .005$, $d = 0.66$ (M$_{MA}$= .757, M$_A$ = .799, SD$_{MA}$

518 = .108, SD$_A$ = .0924). This difference, however, was not reflected in memory bias since we

519 did not find significant differences between the Encoded as A and Encoded as MA sounds in

520 the 2T sequences, where both of the test sounds were presented at encoding, $t(22) = 1.14$, p

521 = .267 (M$_{MA}$ = .509, M$_A$ = .491, SD$_{MA}$ = .0395, SD$_A$ = .0395). The absence of significant

522 differences in memory bias may suggest that they remembered both sounds as evident by the

523 generally high accuracy (i.e., mean performance in the 1T sequences = 0.78 with standard

524 deviation of 0.1) which led them to choose randomly between A and MA sounds in 2T

525 sequences. We complemented the frequentist *t*-tests with corresponding Bayesian *t*-tests,

526 separately for memory performance (1T sequences) and memory bias (2T sequences). The

527 Bayesian *t*-tests for the 1T and 2T sequences yielded similar results as the ones obtained from

528 the frequentist *t*-tests. Specifically, this analysis brought strong evidence for the alternative

529 hypothesis in the case of 1T sequences (BF_{10} = 9.375), while the Bayesian *t*-test for the 2T

530 sequences, brought weak evidence for the alternative hypothesis (BF_{10} = 0.389).
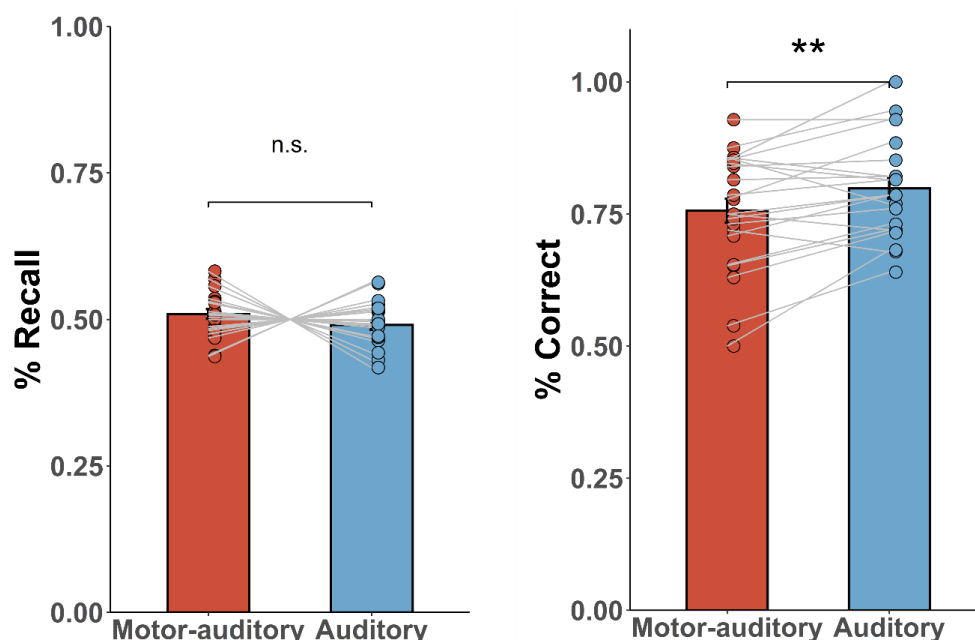
Figure 2. Summary of the behavioural results, separately for memory bias in the 2T sequences (left) and memory performance in the 1T sequences (right). Error bars depict standard error of the mean. For memory bias (i.e., percent recall in 2T sequences), there were no significant differences between motor-auditory and auditory-only sounds (two-tailed paired samples t-test, $p > .050$, $M_{MA} = .509$, $M_A = .491$, $SD_{MA} = .0395$, $SD_A = .0395$), in line with the Bayesian analysis that provided weak evidence for the alternative hypothesis ($BF_{10} = 0.389$). For memory performance (i.e., percent correct in 1T sequences), there was a significant difference between motor-auditory and auditory-only sounds (two-tailed paired samples t-test, $t(22) = 3.15$, $p = .005$, $d = 0.66$), with higher accuracy for the latter ($M_{MA} = .757$, $M_A = .799$, $SD_{MA} = .108$, $SD_A = .0924$), which was also supported by the Bayesian analysis that brought strong evidence in favor of the alternative hypothesis ($BF_{10} = 9.375$).

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 Methods). The time windows defined for each peak were the following: Na 72–92 ms, Tb 120–150 ms, N1/N1$_{mast}$ 94–120 ms, P2/ P2$_{mast}$ 174–224 ms, P3 256–286.

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

19

555 that were passively presented to the participants (Figure 3a-b, see Table 1 for all the mean

556 amplitudes per condition). We also tested for differences in N1 (with reversed polarity) at

557 the mastoids ($N1_{mast}$) using a repeated measures ANOVA with factors Sound Type (MA vs.

558 A) and Laterality (M1 vs. M2). We obtained a significant enhancement for the MA sounds

559 $F(1, 22) = 15.68$, $p < .001$, $\eta_p^2 = .42$, suggesting that besides the attenuation for MA sounds

560 observed at vertex, further modulatory effects of sound-action coincidence occur (Figure 3).

561 We also found a significant main effect of Laterality, $F(1, 22) = 5.96$, $p = .023$, $\eta_p^2 = .21$, with

562 lower amplitudes at M1 compared to M2, while the interaction between Sound Type and

563 Laterality did not reach significance, $F(1, 22) = 3.55$, $p = .073$.

564

565 **Table 1**. Mean amplitudes and standard deviation per component and condition across 23
566 participants.

Components	Electrodes	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	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	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	-	-	-	-	-	-	-	-
$N1_{mast}$	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
$P2_{mast}$	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

567

568 Next, we examined the attenuation effects at the N1 subcomponents at temporal sites,

569 using a 2x2 ANOVA, with factors Sound Type (A vs. MA) and Laterality (T7 vs. T8) on

570 Na and Tb (Figure 3a). For Na, only a significant main effect of Laterality was obtained,

571 with lower amplitudes at T8 compared to T7, $F(1, 22) = 4.82$, $p = .039$, $\eta_p^2 = .18$, while the

572 main effect of Sound Type and the interaction did not reach significance, $F(1, 22) = 0.05$, p

20

573 = .828 and $F(1, 22) = 0.35$, $p = .563$, respectively. For Tb, however, we obtained

574 significantly lower amplitudes for sounds coinciding with a motor act compared to the

575 auditory-only ones, $F(1, 22) = 9.03$, $p = .007$, $\eta_p^2 = .29$, while the main effect of Laterality

576 did not reach significance, $F(1, 22) = 0.03$, $p = .871$. However, we also found a significant

577 interaction, $F(1, 22) = 8.63$, $p = .008$, $\eta_p^2 = .28$, reflecting that the attenuation for MA

578 sounds was only significant in T8 but not in T7 (post-hoc t-tests, $t(22) = -4.06$, $p < .001$, $d =$

579 -0.85 and $t(22) = -1.04$, $p = .311$, respectively).

580   Subsequently, we performed a one-sided t-test to test for possible differences in P2

581 amplitudes between A and MA sounds at encoding, with the hypothesis of attenuated

582 responses for the latter. We obtained a significant P2 attenuation at Cz, $t(22) = 3.98$, $p <$

583 $.001$, $d = 0.83$, with lower amplitudes for sounds that coincided with a motor act, compared

584 to those that were passively presented to the participants (Figure 3a-b). We also tested for

585 differences in this component (with reversed polarity) at the mastoids (P2$_{mast}$) using a

586 repeated measures ANOVA with factors Sound Type (MA vs. A) and Laterality (M1 vs.

587 M2). We observed a significant attenuation for the MA sounds, replicating the attenuation

588 observed at Cz, $F(1, 22) = 34.23$, $p < .001$, $\eta_p^2 = .61$, as well as a main effect of Laterality,

589 $F(1, 22) = 4.66$, $p = .042$, $\eta_p^2 = .17$, with more negative amplitudes at M1 compared to M2.

590 The interaction of Sound Type and Laterality on P2$_{mast}$ did not reach significance, $F(1, 22)$

591 $= 0.54$, $p = .470$. Finally, we also tested for differences in P3 at Pz, which yielded a

592 significantly larger P3 amplitude for sounds coinciding with a motor act, $t(22) = -6.57$, $p <$

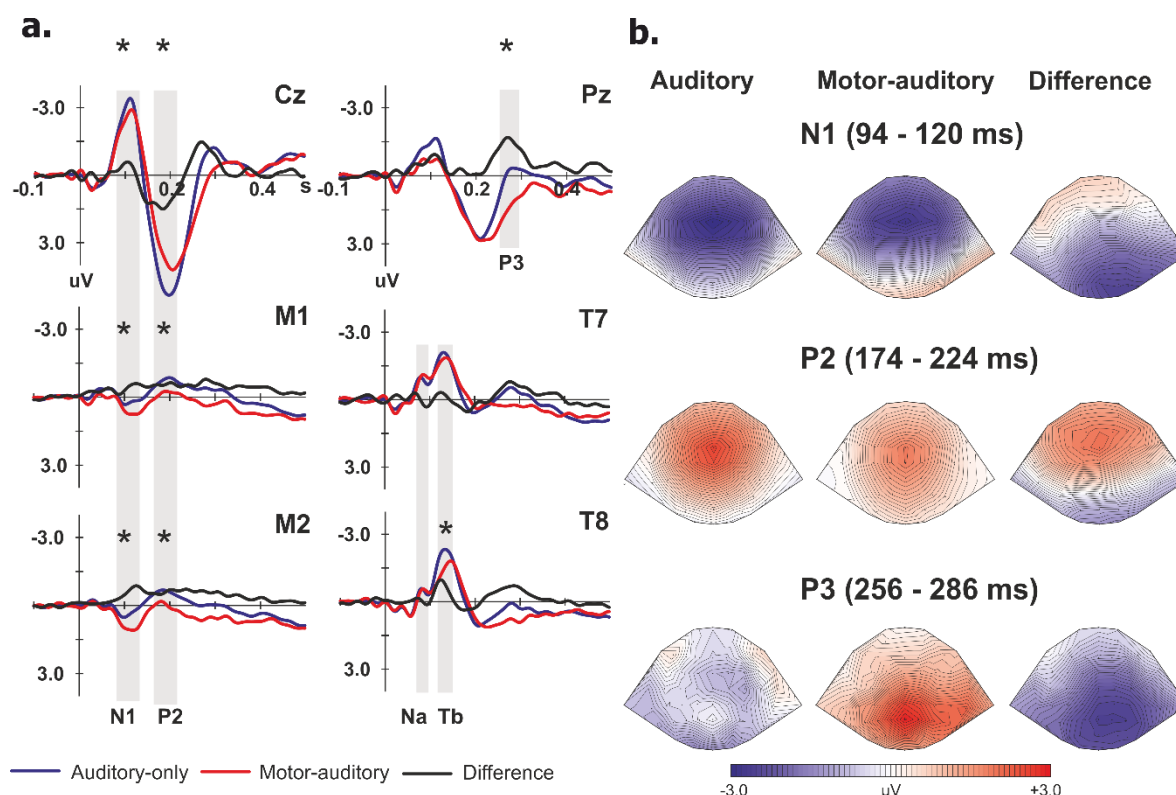593 $.001$, $d = -1.37$ (Figure 3).

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.

3.3. Electrophysiological responses at retrieval

Next, we subdivided the retrieval data depending on whether the sound was encoded as A or MA and whether this sound was recalled or not and we assessed 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$_{mast}$, P2/P2$_{mast}$, 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

22

615 72–92 ms, 120–150 ms, 94–120 ms, 174–224 ms, for the Na, Tb, $N1/N1_{mast}$, and $P2/P2_{mast}$,

616 respectively at the relevant electrodes for each peak.

617 We did not observe any significant effects (all $ps > .05$) on the N1 at Cz and $N1_{mast}$.

618 However, significant results were obtained when we analyzed the modulatory effects of

619 Sound Type and Memory on the N1 subcomponents at temporal sites. We obtained a

620 significant main effect of Sound Type on Na, $F(1, 22) = 7.39$, $p = .013$, $\eta_p^2 = .25$, and Tb,

621 $F(1, 22) = 7.28$, $p = .013$, $\eta_p^2 = .25$, reflecting an enhanced amplitude for sounds that were

622 previously encoded as MA. Additionally, we found a significant interaction between Sound

623 Type and Memory on Na, $F(1, 22) = 5.08$, $p = .035$, $\eta_p^2 = .19$, where post-hoc comparisons

624 showed significantly larger Na amplitude for sounds that were Encoded as MA and were

625 remembered compared to sounds that were Encoded as A and were remembered, $t(45) =$

626 3.73, $p < .001$, $d = 0.55$. In contrast, the post-hoc comparisons did not show significant

627 differences for forgotten sounds as a function of how they were encoded, $t(45) = 0.67$, $p =$

628 .504. No significant differences were found between remembered and forgotten sounds that

629 were Encoded as A, $t(45) = -1.34$, $p = .187$, or between remembered and forgotten sounds

630 that were Encoded as MA, $t(45) = 1.64$, $p = .109$. Similarly, we obtained a significant

631 interaction between Sound Type and Memory on Tb, $F(1, 22) = 4.85$, $p = .038$, $\eta_p^2 = .18$.

632 Post-hoc comparisons showed significantly larger Tb amplitude for sounds that were

633 Encoded as MA and were remembered compared to sounds that were Encoded as A and were

634 remembered, $t(45) = 4.31$, $p < .001$, $d = 0.64$, which is in line with the differences we obtained

635 in the Na window. The post-hoc comparisons also showed lower Tb amplitudes for the

636 Encoded as A sounds when they were remembered compared to when they were forgotten,

637 $t(45) = -3.23$, $p = .002$, $d = -0.48$. Nevertheless, no significant differences were observed

638 between remembered and forgotten sounds that were encoded as MA, $t(45) = 0.64$, $p = .523$,

639 or between the Encoded as MA and Encoded as A sounds that were forgotten, $t(45) = 0.47$,

640 $p = .640$. For both Na and Tb, we did not observe any significant main effects of Laterality,

641 nor any significant interactions between Laterality and Sound Type and/or Memory (all $ps >$

642 0.05). Finally, we did not observe any significant effects on P2 at Cz and $P2_{mast}$ (all $ps > .05$),

643 except for a significant main effect of Memory on $P2_{mast}$, $F(1, 22) = 7.65$, $p = .011$, $\eta_p^2 = .26$,

644 that showed lower amplitudes for sounds that were forgotten ($M_{Forgotten} = -2.08$, $M_{Remembered}$

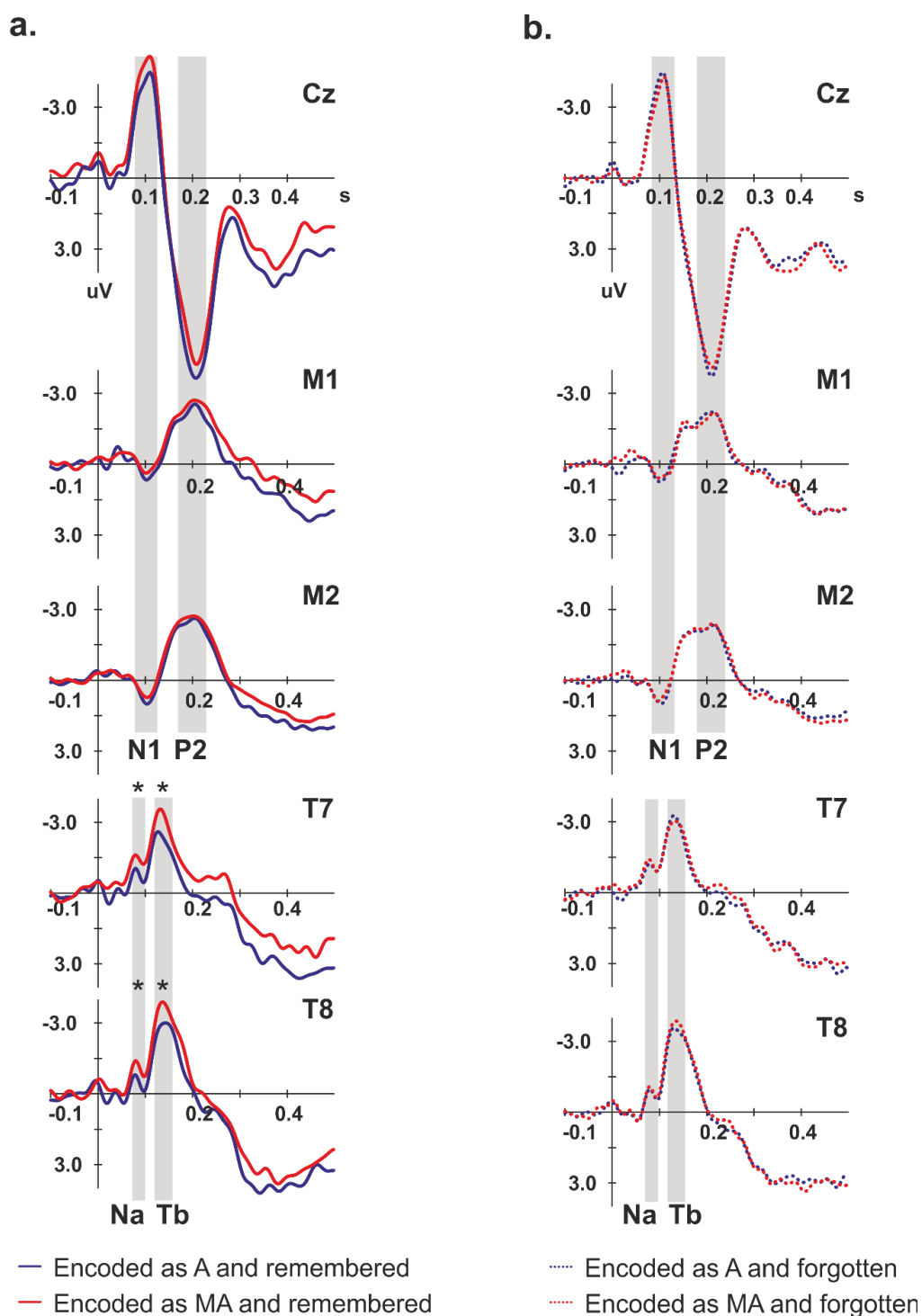645 $= -2.46$, $SD_{Forgotten} = 1.44$, $SD_{Remembered} = 1.56$).

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.

3.3. 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 1,230 ms after sound onset; $p < .05$; Figure 5a) in line with previous animal work (e.g., McGinley et al., 2015). Subsequently, we tested for possible main effects of Sound 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).
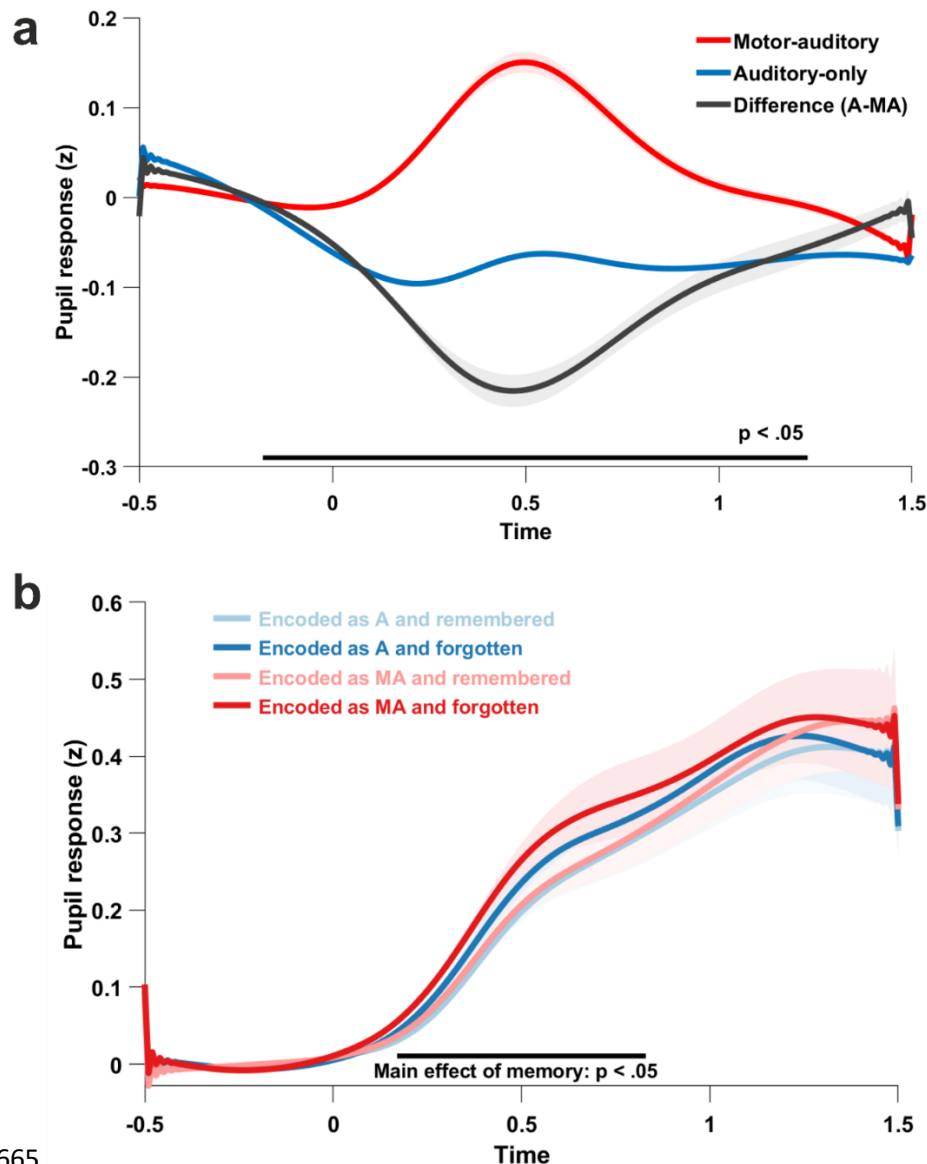
665

Figure 5. **a)** The group-average evoked pupil responses to auditory-only (blue) and motor-auditory (red) events. The effect is depicted as the difference between auditory-only and motor-auditory events (black). Black bar indicates a significant Auditory-only vs. Motor-auditory effect in the window 180 pre-stimulus to 1,230 ms post-stimulus, $p < .05$ (cluster-based permutation test). **b)** The group-average evoked pupil responses 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).

675   **3.4. Correlations**

676   Next, we tested for possible correlations between the behavioural performance,

677   pupillometric and electrophysiological data. For the correlation analyses, we focused on the

678   significant neurophysiological effects at encoding (i.e., ERPs and pupil diameter) and the

679   significant behavioural effect on memory performance. The effects were introduced in the

680   correlation analyses as the difference between A and MA events (see Methods). For the

681   components identified in two electrodes, we calculated the mean amplitude across the two,

682   except for the Tb at encoding, where we introduced only the amplitudes at T8 given the

683   significant interaction between Sound Type and Laterality that showed that attenuation was

684   lateralized. For the pupil data, we calculated the peak of the difference wave (A – MA) within

685   the window of significance (180 ms pre-stimulus until 1,230 ms post-stimulus). All the

686   planned correlations are reported in Table 2.

687

688   **Table 2**. Correlations between the significant self-generation effects. a) electrophysiological
689   effects at encoding (N1, P2, $N1_{mast}$, $P2_{mast}$, P3, and Tb amplitudes) and memory performance
690   (1T sequences), b) neuromodulatory effects at encoding (pupil diameter) and memory
691   performance (1T sequences), c) electrophysiological (N1, P2, $N1_{mast}$, $P2_{mast}$, P3, and Tb
692   amplitudes) and neuromodulatory (pupil diameter) effects at encoding.

693

Correlations between		*r*	*p*
(a) Memory performance	N1	-0.43	**0.041 ***
(1T sequences)	Tb (at T8 only)	-0.55	**0.007****
	P2	-0.19	0.383
	$N1_{mast}$	-0.41	0.055
	$P2_{mast}$	-0.10	0.657
	P3	-0.35	0.098
(b) Memory performance	Pupil diameter	0.46	**0.029***
(1T sequences)			
(c) Pupil diameter	N1	-0.36	0.091
	Tb (at T8 only)	-0.25	0.251
	P2	0.27	0.209
	$N1_{mast}$	-0.23	0.291
	$P2_{mast}$	-0.16	0.507
	P3	-0.08	0.702

694

695   First, we tested whether the significant self-generation effects at encoding (on N1, P2,
696    $N1_{mast}$, $P2_{mast}$, P3, and Tb amplitudes) correlated with the significant self-generation effects
697   on memory performance (1T sequences). This analysis showed a negative correlation
698   between N1 suppression and memory performance (r = -0.43, p = .041; Figure 6a), and a
699   negative correlation between Tb suppression (at T8) and memory performance (r = -0.55, p
700   = .007; Figure 6b), that is, the larger the N1 and Tb suppression, the greater the memory
701   impairment for motor-auditory compared to auditory-only sounds. The remaining
702   correlations did not reach significance (all ps > .05). Second, we assessed whether the
703   difference in pupil diameter between auditory-only and motor-auditory events was related to
704   memory performance and we obtained a significant positive correlation between the two (r
705   = 0.46, p = 0.029; Figure 6c), that is, the larger the pupil dilation for the motor-auditory
706   events, the greater the memory impairment for these sounds. Third, we tested for possible
707   links between the self-generation effects obtained in the ERP analyses (i.e., N1, P2, $N1_{mast}$,
708    $P2_{mast}$, P3 and Tb) and the larger pupil diameter for motor-auditory events. None of these
709   correlations reached significance (all ps > .05), but we observed a non-significant trend
710   towards a correlation between N1 attenuation at Cz and pupil dilation for MA events (Figure
711   6d).

712   Finally, we performed an exploratory correlation analysis to test whether the significant
713   differences in sensory processing we obtained at retrieval between Encoded as A and
714   Encoded as MA sounds were related to the magnitude of the self-generation effects at
715   encoding. To this end, we performed a correlation analysis between the A – MA difference
716   in peaks of the Na and Tb amplitudes (only for the remembered sounds due to the significant
717   interaction) and the effects at encoding (for the N1, P2, $N1_{mast}$, $P2_{mast}$, P3, and Tb amplitudes).
718   We obtained a significant positive correlation between the P2 suppression at encoding and
719   the Na enhancement at retrieval for the remembered sounds, reflecting that the larger the
720   attenuation for P2 at encoding, the larger the Na enhancement for the Encoded as MA sounds
721   that were remembered at retrieval (r = 0.51, p = .012). Similarly, we also obtained a
722   significant negative correlation between Tb at encoding (at T8) and Na for the remembered
723   sounds at retrieval (r = -0.42, p = .04), showing that the larger the attenuation for Tb at
724   encoding, the greater the Na enhancement for motor-auditory sounds that were remembered
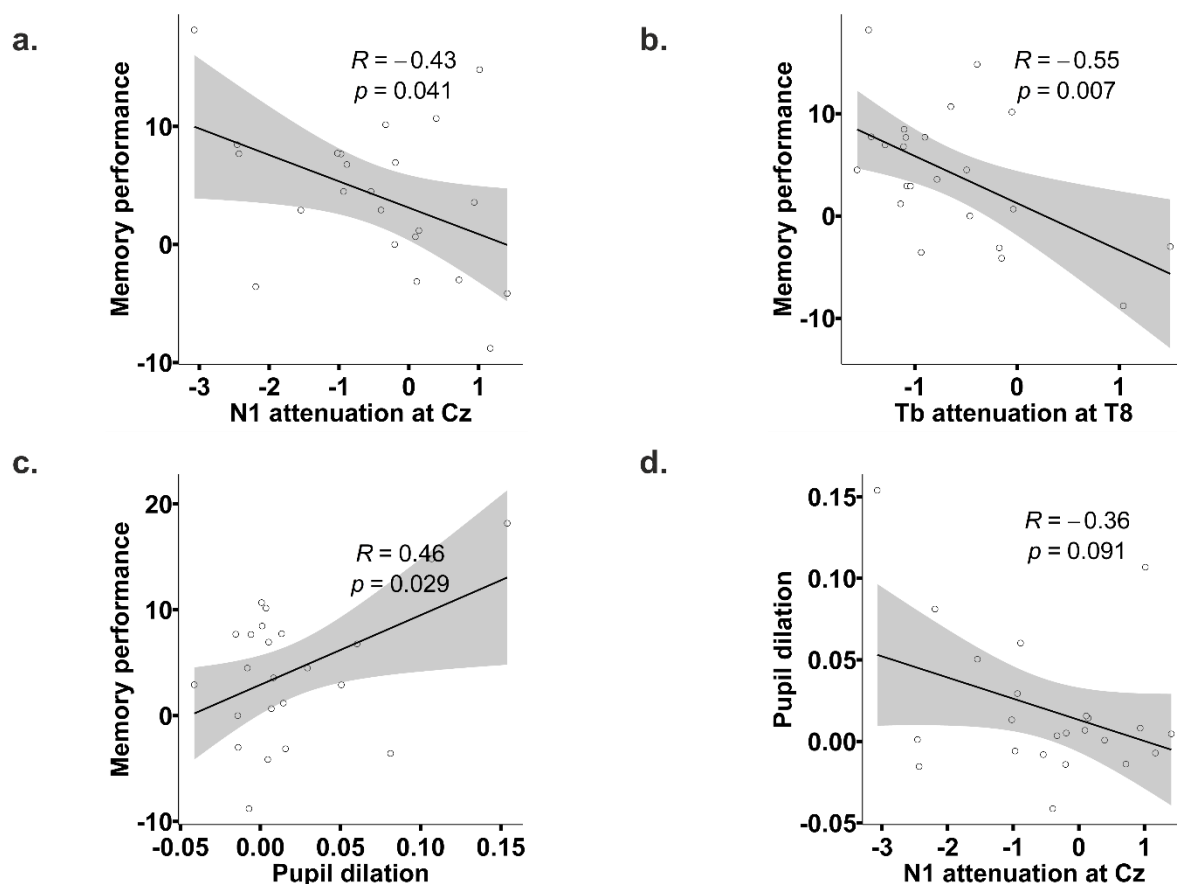725   at retrieval.

Figure 6. Planned correlations between the behavioural, electrophysiological, and pupil data using the Pearson correlation coefficient. **a-b)** Significant negative correlations between N1 suppression (at Cz) and memory performance (r = -0.43, p = .041), and Tb suppression (at T8) and memory performance (r = -0.55, 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 = 0.46, p = 0.029), 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 = -0.36, p = 0.091). The shaded gray areas represent the confidence interval (95% confidence level).

29

4. Discussion

739

740   In this study, we assessed the effects of motor actions on sensory processing and memory

741   encoding of concomitant, but unpredictable sounds, by employing a combination of a self-

742   generation and memory recognition task, while monitoring the brain's and the pupil's

743   responses to sounds that were either presented passively or that coincided in time with a

744   motor act. The aim of the present work was to assess how motor acts affect first sensory

745   processing and second memory encoding of concomitant sounds, and the possible

746   relationships between these two types of effects of actions. Related to the first aim, regarding

747   the effects of actions on sensory processing, we examined whether a) attenuation of sensory

748   processing (i.e., measured by ERPs) prevails even in the absence of a contingent action-sound

749   relationship (e.g., Hórvath et al., 2012), b) actions create a halo of subcortical

750   neuromodulation around them that could be reflected in the pupil diameter (e.g., McGinley

751   et al., 2015), and c) sensory processing (i.e., measured by ERPs) and subcortical

752   neuromodulation (i.e., measured by pupil diameter) were related. Our findings showed N1,

753   P2, $P2_{mast}$, and Tb attenuation for motor-auditory sounds even when they merely coincide

754   with the action, as well as enhancement of P3 and $N1_{mast}$. These findings suggest that self-

755   generation effects are at least partly stimulus-unspecific and driven by alternative

756   mechanisms to the cancellation of predicted sensory reafference via motor forward

757   modelling. Additionally, our data replicated previous animal work (e.g., McGinley et al.,

758   2015) showing that pupil diameter increases dramatically for motor-auditory compared to

759   auditory-only events providing evidence for an alternative stimulus-unspecific mechanism

760   that could underlie sensory suppression for self-generated sounds, namely the activation of

761   subcortical neuromodulation during motor actions. However, the data did not provide clear

762   evidence for a correlation between sensory attenuation and pupil dilation for motor-auditory

763   events. The second aim of the present study was to investigate how actions affect memory

764   encoding of concomitant sounds and whether the potential differences in the memory

765   encoding of motor-auditory and passively presented sounds correlate with sensory

766   suppression and/or subcortical neuromodulation. We found a significant impairment in

767   memory performance for sounds that were encoded as motor-auditory compared to the

768   auditory-only ones demonstrating that the mere presence of an action affects memory

769   encoding of simultaneously presented stimuli. Most importantly, worsened memory

30

770 performance for motor-auditory events correlated with increased sensory suppression (i.e.,

771 N1 and Tb attenuation) and larger pupil dilation for motor-auditory events. These findings

772 fit well with the predictive coding framework suggesting that prediction errors (i.e., reflected

773 in ERPs) drive learning and memory and further support previous work showing that high

774 arousal (i.e., reflected in pupil diameter) may worsen behavioural performance.

775   The first aim of the present study was to assess the effects of actions on auditory

776 processing and subcortical neuromodulation, as well as the relationship between the two.

777 First, we replicate previous work showing that attenuation of N1 and Tb prevails even for

778 mere action-sound coincidences (Horváth et al., 2012, 2013b). Traditionally the N1

779 attenuation has been attributed to predictive processing driven by our actions that attenuate

780 responses in auditory areas when the stimulus can be indeed predicted by the action, implying

781 that attenuation should be specific to the predicted stimulus and thus, mediated by sensory-

782 specific cortices. The stimulus-specificity of the self-generation effects is supported by work

783 showing more pronounced suppression when predictions match more precisely with the

784 sensory input (Fu et al., 2006; Hashimoto & Sakai, 2003; Heinks-Maldonado, Mathalon,

785 Gray, & Ford, 2005; Houde, Nagarajan, Sekihara, & Merzenich, 2002; Baess et al., 2008).

786 However, attenuation of auditory responses occurs also for stimuli merely coinciding with

787 finger movements (Hazemann, Audin, & Lille, 1975; Horváth et al., 2012; Makeig, Muller,

788 & Rockstroh, 1996; Tapia, Cohen, & Starr, 1987) or for unrelated auditory inputs during

789 speech (Numminen, Salmelin, & Hari, 1999), reminiscent of the generalized attenuation

790 found in other sensory modalities during movements (e.g., saccadic suppression or

791 somatosensory gating; Crapse & Sommer, 2008; Williams, Shenasa, & Chapman, 1998;

792 Ross, Morrone, Goldberg, & Burr, 2001). The stimulus-unspecificity of the effects is at least

793 partly supported by evidence suggesting that the N1 and Tb attenuation is probably driven

794 by mere temporal contiguity (Horváth et al., 2012; Hazemann et al., 1975; Han et al., 2021)

795 or temporal predictability (Schafer & Marcus, 1973; Lubinus et al., 2020; Kaiser & Schütz-

796 Bosbach, 2018) and that it mostly reflects modulations of the unspecific component of the

797 auditory N1 (SanMiguel et al., 2013). In line with this evidence, our findings partly point to

798 the modulation of the unspecific component, since for the suppression to be specific to the

799 auditory cortex, N1 should be suppressed at vertex but also at the mastoids. Instead, here we

800 found N1 suppression at vertex, but enhancement on the mastoids for sounds coinciding with

801 actions. Therefore, the effects cannot be entirely specific and action-related activity
802 enhancements – possibly of attentional origin – in auditory areas may take place
803 concomitantly (Horváth, 2015; Schröger et al., 2015; Flinker et al., 2010).

804   The effects on N1 were followed by attenuated P2 and enhanced P3 responses for the
805 sounds coinciding with actions. Although a functional interpretation of P2 is missing
806 (Crowley & Colrain, 2004), empirical evidence has shown that P2 attenuation is mediated by
807 secondary auditory areas (Bosnyak et al., 2004; Pantev et al., 1996), reflecting the processing
808 of the specific features of auditory stimuli (Shahin et al., 2005), and it correlates with the
809 sense of agency (i.e., the feeling of control over actions and their consequences; Gallagher,
810 2000) contrary to the N1 that does not (Ford et al., 2013; Kühn et al., 2011; Timm et al.,
811 2016). These characteristics along with our data showing P2 attenuation in both vertex and
812 mastoids may point to a functional dissociation between N1 and P2 as suggested by previous
813 work (Knolle, Schröger, & Kotz, 2013; Schröger et al., 2015; Chen et al., 2012). Following
814 the P2 attenuation, we found enhanced P3 amplitude at Pz for sounds coinciding with actions.
815 Interestingly, a P3 effect was also evident – although not discussed – in previous work with
816 action-sound coincidences (Horváth et al., 2012). Recently, this effect has been suggested to
817 reflect violations in action-related predictions (Darriba et al., 2021) which may occur in tasks
818 where the self-generated sound is unexpected (e.g., in coincidence tasks where the action
819 does not always result in a sound; Horváth et al., 2012). Although previous work has already
820 described P3 modulations in self-generation paradigms, the posterior distribution and later
821 peak of our effect differentiates it from the fronto-central P3a effect reported for unexpected
822 externally-generated sounds (Baess et al., 2011) or self-generated deviant sounds (Knolle et
823 al., 2013b). Based on previous theories, we speculate that the posterior P3 effect may be
824 related to context updating (Donchin & Coles, 1988), event categorization (Kok, 2001) or
825 decision making (Twomey et al., 2015) and may reflect an evaluative process of the stimulus
826 (i.e., self/external categorization) that ultimately updates the internal model about the sensory
827 consequences of the button press (Polich, 2007).

828   The present study demonstrates that neuromodulatory processes take place concomitantly
829 to the modulatory effects of action-sound coincidence on evoked electrophysiological
830 responses. We obtained pupil dilation measures that are known to track the activity of the
831 LC-NE system (Aston-Jones & Cohen, 2005, Murphy, O'Connell et al., 2014; Joshi et al.,

832   2016) and we showed a remarkable increase in pupil diameter for the motor-auditory events

833   that started even before the action (cf. McGinley et al., 2015a), supporting previous work

834   reporting pupil dilation during finger movements (Lubinus et al., 2021; Yebra et al., 2019),

835   and locomotion (Reimer et al., 2014; Vinck et al., 2015; McGinley et al., 2015) even in the

836   absence of visual stimulation (Richer & Beatty, 1985; Hupe et al., 2009). We hypothesized

837   that these neuromodulatory processes might be behind the stimulus-unspecific effects of

838   actions on simultaneously presented stimuli. However, pupil dilation did not correlate with

839   the sensory suppression effects for self-generated sounds. Although this may suggest that

840   motor-induced sensory suppression and arousal-related neuromodulation during actions

841   operate independently, there was a non-significant trend towards a link between N1

842   attenuation at the vertex and pupil dilation, and both of these measures correlated

843   significantly with memory performance. These inconclusive findings raise the need of future

844   work to further test for relationships between action-induced suppression effects and

845   neuromodulatory mechanisms operating during movement. In fact, previous work has shown

846   that movement correlates with some effects attributed to arousal (Reimer et al., 2014; Vinck

847   et al., 2015), yet, motor signals and arousal-related neuromodulatory inputs have been

848   suggested to exert distinct influences on sensory processing (e.g., McGinley et al., 2015a;

849   Nelson & Mooney, 2016; Reimer et al., 2014; Vinck et al., 2015; for a review see Ferguson

850   & Cardin, 2020). Although we cannot exclude the contribution of other mechanisms, our

851   findings suggest that sensory suppression is not driven by noradrenergic-mediated

852   modulations that have been mainly observed in the sensory thalamus rather than the sensory-

853   specific areas (McBurney et al., 2019), supporting the idea of noradrenergic release from the

854   LC-NE as a "hub" switch for triggering temporally specific, but spatially widespread changes

855   throughout the entire cortex (Kim et al., 2016; Aston-Jones & Cohen, 2005).

856   The second aim of the present study was to assess how the differential processing for

857   sounds coinciding with actions might affect their encoding in memory. While the links

858   between sensorimotor processing of auditory stimuli and memory processes remain largely

859   unexplored, there is evidence that actions attenuate responses in areas supporting memory

860   processes (i.e., Rummell et al., 2016; Mukamel et al., 2010), raising the possibility of a link

861   between self-generation and memory. In our study, motor actions affected the memory

862   encoding of concurrent sounds, but the effects were reflected only in memory performance

863   and not in memory bias. The null effects on memory bias might suggest that participants

864   could recognize that both test sounds at retrieval were presented before, which is supported

865   by the general high level of objective accuracy as well as by reports during an informal

866   debriefing suggesting that many participants thought that most times all sounds at retrieval

867   were presented before. The memory benefit for the more surprising externally-generated

868   sounds fits well with predictive coding theories postulating that items eliciting larger

869   prediction errors at encoding will be encoded better in memory (Henson & Gagnepain, 2010;

870   Greve et al., 2017, 2019; Heilbron & Chait, 2018; Henson & Gagnepain, 2010; Pine et al.,

871   2018; Krawczyk et al., 2018; Rescorla & Wagner,1972; Exton-McGuinness et al., 2015).

872   Yet, one would expect to observe this effect only in contingent paradigms where self-

873   generated sounds are inherently more predictable than the externally-generated ones.

874   However, although in our study actions were not predictive of sound identity or occurrence,

875   they afforded better temporal predictability, which may have compromised the memory

876   encoding of motor-auditory sounds. In fact, the present work is the first human study – to our

877   knowledge – to show that the self-generation effects (i.e., N1 and Tb attenuation) are related

878   to the performance decrements for sounds produced by actions as suggested by previous

879   animal work (Schneider et al., 2018; McGinley et al., 2015; Schneider et al., 2020). These

880   findings support the idea that the larger prediction error responses to unexpected items (as

881   indexed by enhanced ERPs to A compared to MA events at encoding) initiate a cascade of

882   synaptic changes, allowing for more distinctive representations at encoding (Kirwan & Stark,

883   2007, Norman, 2010) and thus better recollection at retrieval. Consistent with this

884   framework, the hippocampus has been implicated as a novelty and match/mismatch detector

885   (Knight, 1996, Stern et al., 1996, Li et al., 2003; Duncan et al., 2012) and there is compelling

886   evidence for hippocampal involvement in learning from prediction errors (Schiffer et al.,

887   2012) and expecting upcoming events (Davachi & DuBrow, 2015, Hindy et al.,

888   2016, Schapiro et al., 2017). Collectively, our data supports the predictive account of

889   memory by showing that sensory attenuation, interpreted as reduced prediction error, is

890   related to decreased memory performance.

891   Memory performance correlated with pupil diameter as well, such that the larger the pupil

892   diameter for motor-auditory events the worse the memory performance for these sounds at

893   retrieval. To date, there have been no direct attempts to test for possible links between motor-

894 induced pupil dilation and memory performance for stimuli triggered by actions. Some

895 interim evidence points to a negative relationship between pupil dilation and detection

896 performance during (McGinley et al., 2015), but also without (Murphy, Vandekerckhove, &

897 Nieuwenhuis, 2014) locomotion, suggesting that performance may follow the classically

898 described, inverted U-shaped dependence on arousal (Yerkes & Dodson, 1908): Intermediate

899 levels of arousal – as indexed by pupil diameter – occur in states of quiet wakefulness and

900 are characterized by optimal performance. In contrast, performance during high-arousal

901 states such as movement, or during quiescence, drops dramatically. Therefore, the observed

902 link between pupil dilation and memory performance in our study may provide yet another

903 piece of evidence supporting the detrimental effects of high arousal on behavioural

904 performance.

905   The present study had clear hypotheses about the effects of actions on sensory and pupil

906 responses at encoding, yet, exploratory analyses of the retrieval data revealed further effects.

907 First, we obtained higher Na and Tb amplitudes for the sounds encoded as motor-auditory

908 and remembered compared to the remembered and encoded as auditory-only ones. Since the

909 sounds encoded as motor-auditory were presented passively at retrieval (i.e., without the

910 motor representation that they were encoded with), the higher Na and Tb amplitudes may

911 reflect a form of contextual prediction error (Exton-McGuinness et al. 2015; Kim et al., 2014;

912 Sinclair & Barense, 2019) due to the mismatch between encoding and retrieval contexts for

913 these sounds. This interpretation can be partly supported by the exploratory correlation

914 analyses that showed that the larger the P2 and Tb attenuation for motor-auditory sounds at

915 encoding, the greater the Na enhancement for these sounds at retrieval when they were

916 remembered. Thus, the greater the effect of the action at encoding, the greater the contextual

917 prediction error when the sound is presented without the action at retrieval. Second, we found

918 larger pupil responses for the forgotten compared to the remembered sounds at retrieval

919 irrespective of how they were encoded. While previous work has reported an old/new pupil

920 effect (i.e., increased pupil responses for the remembered items; Kafkas & Montaldi, 2015;

921 Naber et al., 2013, but see Beukema et al., 2019 for the opposite effect), in our study both

922 sounds at retrieval were presented before. The increase in pupil diameter for the forgotten

923 sounds at retrieval could be instead related to selection or decision uncertainty (Geng et al.,

924   2015; Richer & Beatty, 1987; Nassar et al., 2012; Preuschoff et al., 2011) when participants

925   experienced greater difficulty to decide whether a given sound was presented before or not.

926   In sum, the overarching aim of the present study was to investigate how motor acts affect

927   both sensory processing and the memory encoding of concomitant sounds. To the best of our

928   knowledge, there have been no previous attempts to simultaneously assess the specificity of

929   the self-generation effects and their possible link with neuromodulatory processes while also

930   looking into their possible effects on memory encoding. Here, in a combination of self-

931   generation and memory task, we show that actions affect auditory responses, pupil diameter,

932   and memory encoding of sounds. Actions suppressed sensory responses for concomitant

933   sounds and increased pupil diameter, but these effects were not related, pointing to

934   simultaneous, but probably independent processes. However, sensory suppression and pupil

935   dilation both correlated with memory performance independently, such that the memory

936   performance for sounds coinciding with actions decreased with larger sensory attenuation

937   and greater pupil dilation. Collectively, our findings show self-generation effects even in the

938   absence of a predictive action-sound relationship, replicate previous work showing that pupil

939   diameter increases during actions, and finally point to differentiated internal memory

940   representations for stimuli triggered by ourselves compared to externally presented ones.

941   More importantly, the present study shows that subcortical neuromodulatory systems, along

942   with cortical processes, simultaneously orchestrate auditory processing and memory

943   encoding.

944 **CRediT authorship contribution statement**

945 **Nadia Paraskevoudi:** Conceptualization, Formal Analysis, Methodology, Software,

946 Investigation, Writing - Original draft preparation, Visualization. **Iria SanMiguel:**

947 Conceptualization, Formal Analysis, Methodology, Software, Writing - Original draft

948 preparation, Supervision, Project administration, Funding acquisition.

949

960

961

962   **References**

963   Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-

964   norepinephrine function: Adaptive Gain and Optimal Performance. *Annual Review*

965   *of Neuroscience*, *28*(1), 403–450.

966   https://doi.org/10.1146/annurev.neuro.28.061604.135709

967   Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-

968   initiated sounds in an auditory stream: An ERP study: Selective suppression of self-

969   initiated sounds. *Psychophysiology*, *48*(9), 1276–1283.

970   https://doi.org/10.1111/j.1469-8986.2011.01196.x

971   Bar, M. (2009). The proactive brain: Memory for predictions. *Philosophical Transactions*

972   *of the Royal Society B: Biological Sciences*, *364*(1521), 1235–1243.

973   https://doi.org/10.1098/rstb.2008.0310

974   Bäß, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related

975   potential component with unpredictable self-initiated tones: Evidence for internal

976   forward models with dynamic stimulation. *International Journal of*

977   *Psychophysiology*, *70*(2), 137–143. https://doi.org/10.1016/j.ijpsycho.2008.06.005

978   Beukema, S., Jennings, B. J., Olson, J. A., & Kingdom, F. A. A. (2019). The Pupillary

979   Response to the Unknown: Novelty Versus Familiarity. *I-Perception*, *10*(5),

980   204166951987481. https://doi.org/10.1177/2041669519874817

981   Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to Bright Surfaces Enhances

982   the Pupillary Light Reflex. *Journal of Neuroscience*, *33*(5), 2199–2204.

983   https://doi.org/10.1523/JNEUROSCI.3440-12.2013

984   Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-

985   produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640.

986   https://doi.org/10.1038/2870

987   Bosnyak, D. J. (2004). Distributed Auditory Cortical Representations Are Modified When

988   Non-musicians Are Trained at Pitch Discrimination with 40 Hz Amplitude

989   Modulated Tones. *Cerebral Cortex*, *14*(10), 1088–1099.

990   https://doi.org/10.1093/cercor/bhh068

991   Brown, R. M., & Palmer, C. (2012). Auditory–motor learning influences auditory memory

992   for music. *Memory & Cognition*, *40*(4), 567–578. https://doi.org/10.3758/s13421-

993   011-0177-x

994   Chagnaud, B. P., Banchi, R., Simmers, J., & Straka, H. (2015). Spinal corollary discharge

995   modulates motion sensing during vertebrate locomotion. *Nature Communications*,

996   *6*(1), 7982. https://doi.org/10.1038/ncomms8982

997   Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten Percent Electrode System for

998   Topographic Studies of Spontaneous and Evoked EEG Activities. *American*

999   *Journal of EEG Technology*, *25*(2), 83–92.

1000   https://doi.org/10.1080/00029238.1985.11080163

1001   Chen, Z., Chen, X., Liu, P., Huang, D., & Liu, H. (2012). Effect of temporal predictability

1002   on the neural processing of self-triggered auditory stimulation during vocalization.

1003   *BMC Neuroscience*, *13*(1), 55. https://doi.org/10.1186/1471-2202-13-55

1004   Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of*

1005   *Memory and Language*, *26*(3), 341–361. https://doi.org/10.1016/0749-

1006   596X(87)90118-5

1007   Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom.

1008   *Nature Reviews Neuroscience*, *9*(8), 587–600. https://doi.org/10.1038/nrn2457

1009   Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an

1010   independent component process: Age, sleep and modality. *Clinical*

1011   *Neurophysiology*, *115*(4), 732–744. https://doi.org/10.1016/j.clinph.2003.11.021

1012   Darriba, Á., Hsu, Y.-F., Van Ommen, S., & Waszak, F. (2021). Intention-based and

1013   sensory-based predictions. *Scientific Reports*, *11*(1), 19899.

1014   https://doi.org/10.1038/s41598-021-99445-z

1015   Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of

1016   prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99.

1017   https://doi.org/10.1016/j.tics.2014.12.004

1018   Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of

1019   single-trial EEG dynamics including independent component analysis. *Journal of*

1020   *Neuroscience Methods*, *134*(1), 9–21.

1021   https://doi.org/10.1016/j.jneumeth.2003.10.009

1022   Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context

1023   updating? *Behavioral and Brain Sciences*, *11*(03), 357.

1024   https://doi.org/10.1017/S0140525X00058027

1025   Duncan, K., Ketz, N., Inati, S. J., & Davachi, L. (2012). Evidence for area CA1 as a

1026   match/mismatch detector: A high-resolution fMRI study of the human

1027   hippocampus. *Hippocampus*, *22*(3), 389–398. https://doi.org/10.1002/hipo.20933

1028   Exton-McGuinness, M. T. J., Lee, J. L. C., & Reichelt, A. C. (2015). Updating memories—

1029   The role of prediction errors in memory reconsolidation. *Behavioural Brain*

1030   *Research*, *278*, 375–384. https://doi.org/10.1016/j.bbr.2014.10.011

1031   Ferguson, K. A., & Cardin, J. A. (2020). Mechanisms underlying gain modulation in the

1032   cortex. *Nature Reviews Neuroscience*, *21*(2), 80–92.

1033   https://doi.org/10.1038/s41583-019-0253-y

1034   Flinker, A., Chang, E. F., Kirsch, H. E., Barbaro, N. M., Crone, N. E., & Knight, R. T.

1035   (2010). Single-Trial Speech Suppression of Auditory Cortex Activity in Humans.

1036   *Journal of Neuroscience*, *30*(49), 16643–16650.

1037   https://doi.org/10.1523/JNEUROSCI.1809-10.2010

1038   Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2014). Did I Do That?

1039   Abnormal Predictive Processes in Schizophrenia When Button Pressing to Deliver a

1040   Tone. *Schizophrenia Bulletin*, *40*(4), 804–812.

1041   https://doi.org/10.1093/schbul/sbt072

1042   Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal*

1043   *Society B: Biological Sciences*, *360*(1456), 815–836.

1044   https://doi.org/10.1098/rstb.2005.1622

1045   Fu, C. H. Y., Vythelingum, G. N., Brammer, M. J., Williams, S. C. R., Amaro, E., Andrew,

1046   C. M., Yágüez, L., van Haren, N. E. M., Matsumoto, K., & McGuire, P. K. (2006).

1047   An fMRI Study of Verbal Self-monitoring: Neural Correlates of Auditory Verbal

1048   Feedback. *Cerebral Cortex*, *16*(7), 969–977. https://doi.org/10.1093/cercor/bhj039

1049   Gagnepain, P., Henson, R., Chételat, G., Desgranges, B., Lebreton, K., & Eustache, F.

1050   (2011). Is Neocortical–Hippocampal Connectivity a Better Predictor of Subsequent

1051   Recollection than Local Increases in Hippocampal Activity? New Insights on the

1052   Role of Priming. *Journal of Cognitive Neuroscience*, *23*(2), 391–403.

1053   https://doi.org/10.1162/jocn.2010.21454

1054   Geng, J. J., Blumenfeld, Z., Tyson, T. L., & Minzenberg, M. J. (2015). Pupil diameter

1055   reflects uncertainty in attentional selection during visual search. *Frontiers in Human*

1056   *Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00435

1057   Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction

1058   error drive one-shot declarative learning? *Journal of Memory and Language*, *94*,

1059   149–165. https://doi.org/10.1016/j.jml.2016.11.001

1060   Halgren, E. (1991). Firing of human hippocampal units in relation to voluntary movements.

1061   *Hippocampus*, *1*(2), 153–161. https://doi.org/10.1002/hipo.450010204

1062   Han, N., Jack, B. N., Hughes, G., Elijah, R. B., & Whitford, T. J. (2021). Sensory

1063   attenuation in the absence of movement: Differentiating motor action from sense of

1064   agency. *Cortex*, *141*, 436–448. https://doi.org/10.1016/j.cortex.2021.04.010

1065   Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring

1066   of speech production with delayed auditory feedback: An fMRI study. *Human*

1067   *Brain Mapping*, *20*(1), 22–28. https://doi.org/10.1002/hbm.10119

1068   Hazemann, P., Audin, G., & Lille, F. (1975). Effect of voluntary self-paced movements

1069   upon auditory and somatosensory evoked potentials in man.

1070   *Electroencephalography and Clinical Neurophysiology*, *39*(3), 247–254.

1071   https://doi.org/10.1016/0013-4694(75)90146-7

1072   Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive

1073   Coding in Auditory Cortex? *Neuroscience*, *389*, 54–73.

1074   https://doi.org/10.1016/j.neuroscience.2017.07.061

1075   Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of

1076   auditory cortex during speech production. *Psychophysiology*, *42*(2), 180–190.

1077   https://doi.org/10.1111/j.1469-8986.2005.00272.x

1078   Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems.

1079   *Hippocampus*, *20*(11), 1315–1326. https://doi.org/10.1002/hipo.20857

1080   Hesse, M. D., Nishitani, N., Fink, G. R., Jousmaki, V., & Hari, R. (2010). Attenuation of

1081   Somatosensory Responses to Self-Produced Tactile Stimulation. *Cerebral Cortex*,

1082   *20*(2), 425–432. https://doi.org/10.1093/cercor/bhp110

1083   Hindy, N. C., Ng, F. Y., & Turk-Browne, N. B. (2016). Linking pattern completion in the

1084   hippocampus to predictive coding in visual cortex. *Nature Neuroscience*, *19*(5),

1085   665–667. https://doi.org/10.1038/nn.4284

1086   Hoeks, B., & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: A

1087   quantitative system analysis. *Behavior Research Methods, Instruments, &*

1088   *Computers*, *25*(1), 16–26. https://doi.org/10.3758/BF03204445

1089   Horváth, J. (2013a). Attenuation of auditory ERPs to action-sound coincidences is not

1090   explained by voluntary allocation of attention: Action-sound coincidence effect is

1091   not attentional. *Psychophysiology*, *50*(3), 266–273.

1092   https://doi.org/10.1111/psyp.12009

1093   Horváth, J. (2013b). Action-sound coincidence-related attenuation of auditory ERPs is not

1094   modulated by affordance compatibility. *Biological Psychology*, *93*(1), 81–87.

1095   https://doi.org/10.1016/j.biopsycho.2012.12.008

1096   Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses.

1097   *Brain Research*, *1626*, 54–65. https://doi.org/10.1016/j.brainres.2015.03.038

1098   Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action–Sound Coincidences Suppress

1099   Evoked Responses of the Human Auditory Cortex in EEG and MEG. *Journal of*

1100   *Cognitive Neuroscience*, *24*(9), 1919–1931. https://doi.org/10.1162/jocn_a_00215

1101   Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the

1102   Auditory Cortex during Speech: An MEG Study. *Journal of Cognitive*

1103   *Neuroscience*, *14*(8), 1125–1138. https://doi.org/10.1162/089892902760807140

1104   Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual

1105   sensory attenuation in voluntary action. *NeuroImage*, *56*(3), 1632–1640.

1106   https://doi.org/10.1016/j.neuroimage.2011.02.057

1107   Hupe, J. M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion

1108   perception. *Journal of Vision*, *9*(7), 10–10. https://doi.org/10.1167/9.7.10

1109   Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil

1110   Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate

1111   Cortex. *Neuron*, *89*(1), 221–234. https://doi.org/10.1016/j.neuron.2015.11.028

1112   Kafkas, A., & Montaldi, D. (2015). The pupillary response discriminates between

1113   subjective and objective familiarity and novelty. *Psychophysiology*, *52*(10), 1305–

1114   1316. https://doi.org/10.1111/psyp.12471

1115   Kaiser, J., & Schütz-Bosbach, S. (2018). Sensory attenuation of self-produced signals does

1116   not rely on self-specific motor predictions. *European Journal of Neuroscience*,

1117   *47*(11), 1303–1310. https://doi.org/10.1111/ejn.13931

1118   Kelley, D. B., & Bass, A. H. (2010). Neurobiology of vocal communication: Mechanisms

1119   for sensorimotor integration and vocal patterning. *Current Opinion in*

1120   *Neurobiology*, *20*(6), 748–753. https://doi.org/10.1016/j.conb.2010.08.007

1121   Kilteni, K., Engeler, P., & Ehrsson, H. H. (2020). Efference Copy Is Necessary for the

1122   Attenuation of Self-Generated Touch. *IScience*, *23*(2), 100843.

1123   https://doi.org/10.1016/j.isci.2020.100843

1124   Kim, A. J., Fitzgerald, J. K., & Maimon, G. (2015). Cellular evidence for efference copy in

1125   Drosophila visuomotor processing. *Nature Neuroscience*, *18*(9), 1247–1255.

1126   https://doi.org/10.1038/nn.4083

1127   Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014). Pruning of

1128   memories by context-based prediction error. *Proceedings of the National Academy*

1129   *of Sciences*, *111*(24), 8997–9002. https://doi.org/10.1073/pnas.1319438111

1130   Kim, J.-H., Jung, A.-H., Jeong, D., Choi, I., Kim, K., Shin, S., Kim, S. J., & Lee, S.-H.

1131   (2016). Selectivity of Neuromodulatory Projections from the Basal Forebrain and

1132   Locus Ceruleus to Primary Sensory Cortices. *The Journal of Neuroscience*, *36*(19),

1133   5314–5327. https://doi.org/10.1523/JNEUROSCI.4333-15.2016

1134   Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation

1135   of pattern separation in the medial temporal lobe. *Learning & Memory*, *14*(9),

1136   625–633. https://doi.org/10.1101/lm.663507

1137   Klaffehn, A. L., Baess, P., Kunde, W., & Pfister, R. (2019). Sensory attenuation prevails

1138   when controlling for temporal predictability of self- and externally generated tones.

1139   *Neuropsychologia*, *132*, 107145.

1140   https://doi.org/10.1016/j.neuropsychologia.2019.107145

1141   Knapen, T., de Gee, J. W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., & Theeuwes, J.

1142   (2016). Cognitive and Ocular Factors Jointly Determine Pupil Responses under

1143   Equiluminance. *PLOS ONE*, *11*(5), e0155574.

1144   https://doi.org/10.1371/journal.pone.0155574

1145   Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection.

1146   *Nature*, *383*(6597), 256–259. https://doi.org/10.1038/383256a0

1147   Knolle, F., Schröger, E., & Kotz, S. A. (2013a). Prediction errors in self- and externally-

1148   generated deviants. *Biological Psychology*, *92*(2), 410–416.

1149   https://doi.org/10.1016/j.biopsycho.2012.11.017

1150   Knolle, F., Schröger, E., & Kotz, S. A. (2013b). Cerebellar contribution to the prediction of

1151   self-initiated sounds. *Cortex*, *49*(9), 2449–2461.

1152   https://doi.org/10.1016/j.cortex.2012.12.012

1153   Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity.

1154   *Psychophysiology*, *38*(3), 557–577. https://doi.org/10.1017/S0048577201990559

1155   Korka, B., Schröger, E., & Widmann, A. (2019). Action Intention-based and Stimulus

1156   Regularity-based Predictions: Same or Different? *Journal of Cognitive*

1157   *Neuroscience*, *31*(12), 1917–1932. https://doi.org/10.1162/jocn_a_01456

1158   Korka, B., Schröger, E., & Widmann, A. (2020). What *exactly* is missing here? The sensory

1159   processing of unpredictable omissions is modulated by the specificity of expected

1160   action-effects. *European Journal of Neuroscience*, *52*(12), 4667–4683.

1161   https://doi.org/10.1111/ejn.14899

1162   Krawczyk, M. C., Fernández, R. S., Pedreira, M. E., & Boccia, M. M. (2017). Toward a

1163   better understanding on the role of prediction error on memory processes: From

1164   bench to clinic. *Neurobiology of Learning and Memory*, *142*, 13–20.

1165   https://doi.org/10.1016/j.nlm.2016.12.011

1166   Kühn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., & Voss, M. (2011). Whodunnit?

1167   Electrophysiological Correlates of Agency Judgements. *PLoS ONE*, *6*(12), e28657.

1168   https://doi.org/10.1371/journal.pone.0028657

1169   Lee, C. R., & Margolis, D. J. (2016). Pupil Dynamics Reflect Behavioral Choice and

1170   Learning in a Go/NoGo Tactile Decision-Making Task in Mice. *Frontiers in*

1171   *Behavioral Neuroscience*, *10*. https://doi.org/10.3389/fnbeh.2016.00200

1172   Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical*

1173   *Course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

1174   Li, S., Cullen, W. K., Anwyl, R., & Rowan, M. J. (2003). Dopamine-dependent facilitation

1175   of LTP induction in hippocampal CA1 by exposure to spatial novelty. *Nature*

1176   *Neuroscience*, *6*(5), 526–531. https://doi.org/10.1038/nn1049

1177   Lubinus, C., Einhäuser, W., Schiller, F., Kircher, T., Straube, B., & van Kemenade, B. M.

1178   (2021). *Action-based predictions affect visual perception, neural processing, and*

pupil size, regardless of temporal predictability [Preprint]. 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). https://doi.org/10.1007/BF00229149

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

Martikainen, M. H. (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

McBurney-Lin, J., Lu, J., Zuo, Y., & Yang, H. (2019). Locus coeruleus-norepinephrine modulation of sensory processing and perception: A focused review. *Neuroscience & Biobehavioral Reviews*, *105*, 190–199. https://doi.org/10.1016/j.neubiorev.2019.06.009

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

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

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*

1210   *Psychologie Expérimentale*, *58*(3), 206–219. https://doi.org/10.1037/h0087445

1211   Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-Neuron

1212   Responses in Humans during Execution and Observation of Actions. *Current*

1213   *Biology*, *20*(8), 750–756. https://doi.org/10.1016/j.cub.2010.02.045

1214   Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014).

1215   Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human*

1216   *Brain Mapping*, *35*(8), 4140–4154. https://doi.org/10.1002/hbm.22466

1217   Murphy, P. R., Vandekerckhove, J., & Nieuwenhuis, S. (2014). Pupil-Linked Arousal

1218   Determines Variability in Perceptual Decision Making. *PLoS Computational*

1219   *Biology*, *10*(9), e1003854. https://doi.org/10.1371/journal.pcbi.1003854

1220   Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic

1221   Response to Sound: A Review and an Analysis of the Component Structure.

1222   *Psychophysiology*, *24*(4), 375–425. https://doi.org/10.1111/j.1469-

1223   8986.1987.tb00311.x

1224   Naber, M., Frassle, S., Rutishauser, U., & Einhauser, W. (2013). Pupil size signals novelty

1225   and predicts later retrieval success for declarative memories of natural scenes.

1226   *Journal of Vision*, *13*(2), 11–11. https://doi.org/10.1167/13.2.11

1227   Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012).

1228   Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature*

1229   *Neuroscience*, *15*(7), 1040–1046. https://doi.org/10.1038/nn.3130

1230   Nelson, A., & Mooney, R. (2016). The Basal Forebrain and Motor Cortex Provide

1231   Convergent yet Distinct Movement-Related Inputs to the Auditory Cortex. *Neuron*,

1232   *90*(3), 635–648. https://doi.org/10.1016/j.neuron.2016.03.031

1233   Norman, K. A. (2010). How hippocampus and cortex contribute to recognition memory:

1234   Revisiting the complementary learning systems model. *Hippocampus*, *20*(11),

1235   1217–1227. https://doi.org/10.1002/hipo.20855

1236   Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of

1237   the human auditory cortex. *Neuroscience Letters*, *265*(2), 119–122.

1238   https://doi.org/10.1016/S0304-3940(99)00218-9

1239   Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain

1240   dynamics. In *Progress in Brain Research* (Vol. 159, pp. 99–120). Elsevier.

1241   https://doi.org/10.1016/S0079-6123(06)59007-7

1242   Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source

1243   Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological

1244   Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9.

1245   https://doi.org/10.1155/2011/156869

1246   Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-

1247   resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–

1248   719. https://doi.org/10.1016/S1388-2457(00)00527-7

1249   Ozubko, J. D., Gopie, N., & MacLeod, C. M. (2012). Production benefits both recollection

1250   and familiarity. *Memory & Cognition*, *40*(3), 326–338.

1251   https://doi.org/10.3758/s13421-011-0165-1

1252   Pantev, C., Eulitz, C., Hampson, S., Ross, B., & Roberts, L. E. (1996). The Auditory

1253   Evoked "Off" Response: Sources and Comparison with the"On" and the

1254   "Sustained" Responses: *Ear and Hearing*, *17*(3), 255–265.

1255   https://doi.org/10.1097/00003446-199606000-00008

1256   Pine, A., Sadeh, N., Ben-Yakov, A., Dudai, Y., & Mendelsohn, A. (2018). Knowledge

1257   acquisition is governed by striatal prediction errors. *Nature Communications*, *9*(1),

1258   1673. https://doi.org/10.1038/s41467-018-03992-5

1259   Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical*

1260   *Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

1261   Press, C., & Cook, R. (2015). Beyond action-specific simulation: Domain-general motor

1262   contributions to perception. *Trends in Cognitive Sciences*, *19*(4), 176–178.

1263   https://doi.org/10.1016/j.tics.2015.01.006

1264   Press, C., Kok, P., & Yon, D. (2020). The Perceptual Prediction Paradox. *Trends in*

1265   *Cognitive Sciences*, *24*(1), 13–24. https://doi.org/10.1016/j.tics.2019.11.003

1266   Preuschoff, K. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in

1267   decision making. *Frontiers in Neuroscience*, *5*.

1268   https://doi.org/10.3389/fnins.2011.00115

1269   Pyasik, M., Burin, D., & Pia, L. (2018). On the relation between body ownership and sense

1270   of agency: A link at the level of sensory-related signals. *Acta Psychologica*, *185*,

1271   219–228. https://doi.org/10.1016/j.actpsy.2018.03.001

1272   Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S.

1273   (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet

1274   Wakefulness. *Neuron*, *84*(2), 355–362.

1275   https://doi.org/10.1016/j.neuron.2014.09.033

1276   Requarth, T., & Sawtell, N. B. (2011). Neural mechanisms for filtering self-generated

1277   sensory signals in cerebellum-like circuits. *Current Opinion in Neurobiology*, *21*(4),

1278   602–608. https://doi.org/10.1016/j.conb.2011.05.031

1279   Rescorla, R. A., & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations

1280   in the Effectiveness of Reinforcement and Nonreinforcement. In A. H. Black, & W.

1281   F. Prokasy (Eds.), Classical Conditioning II: Current Research and Theory (pp. 64-

1282   99). New York: Appleton- Century-Crofts.

1283   Richer, F., & Beatty, J. (1987). Contrasting Effects of Response Uncertainty on the Task-

1284   Evoked Pupillary Response and Reaction Time. *Psychophysiology*, *24*(3), 258–262.

1285   https://doi.org/10.1111/j.1469-8986.1987.tb00291.x

1286   Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual

1287   perception at the time of saccades. *Trends in Neurosciences*, *24*(2), 113–121.

1288   https://doi.org/10.1016/S0166-2236(00)01685-4

1289   Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes

1290   factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.

1291   https://doi.org/10.1016/j.jmp.2012.08.001

1292   Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory

1293   attenuation. *Neuropsychologia*, *51*(5), 922–929.

1294   https://doi.org/10.1016/j.neuropsychologia.2013.02.005

1295   Roussel, C., Hughes, G., & Waszak, F. (2014). Action prediction modulates both

1296   neurophysiological and psychophysical indices of sensory attenuation. *Frontiers in*

1297   *Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00115

1298   Roy, J. E., & Cullen, K. E. (2001). Selective Processing of Vestibular Reafference during

1299   Self-Generated Head Motion. *The Journal of Neuroscience*, *21*(6), 2131–2142.

1300   https://doi.org/10.1523/JNEUROSCI.21-06-02131.2001

1301   Rummell, B. P., Klee, J. L., & Sigurdsson, T. (2016). Attenuation of Responses to Self-

1302   Generated Sounds in Auditory Cortical Neurons. *The Journal of Neuroscience*,

1303   *36*(47), 12010–12026. https://doi.org/10.1523/JNEUROSCI.1564-16.2016

1304   SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated

1305   sounds reflect the attenuation of the unspecific N1 component of the auditory ERP:

1306   Auditory N1 suppression: N1 components. *Psychophysiology*, *50*(4), 334–343.

1307   https://doi.org/10.1111/psyp.12024

1308   Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial

1309   suppression of self-generated sounds and its dependence on attention. *International*

1310   *Journal of Psychophysiology*, *90*(3), 300–310.

1311   https://doi.org/10.1016/j.ijpsycho.2013.09.006

1312   Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain

1313   Responses. *Science*, *181*(4095), 175–177.

1314   https://doi.org/10.1126/science.181.4095.175

1315   Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017).

1316   Complementary learning systems within the hippocampus: A neural network

1317   modelling approach to reconciling episodic memory with statistical learning.

1318   *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711),

1319   20160049. https://doi.org/10.1098/rstb.2016.0049

1320   Schiffer, A.-M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (2012). Surprised at All the

1321   Entropy: Hippocampal, Caudate and Midbrain Contributions to Learning from

1322   Prediction Errors. *PLoS ONE*, *7*(5), e36445.

1323   https://doi.org/10.1371/journal.pone.0036445

1324   Schneider, D. M. (2020). Reflections of action in sensory cortex. *Current Opinion in*

1325   *Neurobiology*, *64*, 53–59. https://doi.org/10.1016/j.conb.2020.02.004

1326   Schneider, D. M., & Mooney, R. (2018). How Movement Modulates Hearing. *Annual*

1327   *Review of Neuroscience*, *41*(1), 553–572. https://doi.org/10.1146/annurev-neuro-

1328   072116-031215

1329   Schneider, D. M., Nelson, A., & Mooney, R. (2014). A synaptic and circuit basis for

1330   corollary discharge in the auditory cortex. *Nature*, *513*(7517), 189–194.

1331   https://doi.org/10.1038/nature13724

1332   Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human

1333   audition: A lesson from cognitive psychophysiology. *European Journal of*

1334   *Neuroscience*, *41*(5), 641–664. https://doi.org/10.1111/ejn.12816

1335   Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2

1336   auditory-evoked responses by the spectral complexity of musical sounds.

1337   *NeuroReport*, *16*(16), 1781–1785.

1338   https://doi.org/10.1097/01.wnr.0000185017.29316.63

1339   Simpson, H. M. (1969). Effects of a task-relevant response on pupil size.

1340   *Psychophysiology*, *6*(2), 115–121. https://doi.org/10.1111/j.1469-

1341   8986.1969.tb02890.x

1342   Sinclair, A. H., & Barense, M. D. (2019). Prediction Error and Memory Reactivation: How

1343   Incomplete Reminders Drive Reconsolidation. *Trends in Neurosciences*, *42*(10),

1344   727–739. https://doi.org/10.1016/j.tins.2019.08.007

1345   Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by

1346   visual inversion. *Journal of Comparative and Physiological Psychology*, *43*(6),

1347   482–489. https://doi.org/10.1037/h0055479

1348   Stern, C. E., Corkin, S., Gonzalez, R. G., Guimaraes, A. R., Baker, J. R., Jennings, P. J.,

1349   Carr, C. A., Sugiura, R. M., Vedantham, V., & Rosen, B. R. (1996). The

1350   hippocampal formation participates in novel picture encoding: Evidence from

1351   functional magnetic resonance imaging. *Proceedings of the National Academy of*

1352   *Sciences*, *93*(16), 8660–8665. https://doi.org/10.1073/pnas.93.16.8660

1353   Strauch, C., Koniakowsky, I., & Huckauf, A. (2020). Decision Making and Oddball Effects

1354   on Pupil Size: Evidence for a Sequential Process. *Journal of Cognition*, *3*(1), 7.

1355   https://doi.org/10.5334/joc.96

1356   Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D.

1357   (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*,

1358   *364*(6437), eaav7893. https://doi.org/10.1126/science.aav7893

1359   Tapia, M. C., Cohen, L. G., & Starr, A. (1987). Attenuation of Auditory-Evoked Potentials

1360   during Voluntary Movement in Man. *International Journal of Audiology*, *26*(6),

1361   369–373. https://doi.org/10.3109/00206098709081565

1362   Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for

1363   self-initiated sounds is independent of attention. *BMC Neuroscience*, *14*(1), 2.

1364   https://doi.org/10.1186/1471-2202-14-2

Timm, J., Schönwiesner, M., Schröger, E., & SanMiguel, I. (2016). Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. *Cortex*, *80*, 5–20. https://doi.org/10.1016/j.cortex.2016.03.018

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. (1954). Relations between the central Nervous System and the peripheral organs. *The British Journal of Animal Behaviour*, *2*(3), 89–94. https://doi.org/10.1016/S0950-5601(54)80044-X

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

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.

1396   https://doi.org/10.1152/jn.1998.79.2.947

1397   Wolpaw, J. R., & Penry, J. K. (1975). A temporal component of the auditory evoked

1398   response. *Electroencephalography and Clinical Neurophysiology*, *39*(6), 609–620.

1399   https://doi.org/10.1016/0013-4694(75)90073-5

1400   Wolpert, D., Ghahramani, Z., & Jordan, M. (1995). An internal model for sensorimotor

1401   integration. *Science*, *269*(5232), 1880–1882.

1402   https://doi.org/10.1126/science.7569931

1403   Yebra, M., Galarza-Vallejo, A., Soto-Leon, V., Gonzalez-Rosa, J. J., de Berker, A. O.,

1404   Bestmann, S., Oliviero, A., Kroes, M. C. W., & Strange, B. A. (2019). Action

1405   boosts episodic memory encoding in humans via engagement of a noradrenergic

1406   system. *Nature Communications*, *10*(1), 3534. https://doi.org/10.1038/s41467-019-

1407   11358-8

1408   Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of

1409   habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–

1410   482. https://doi.org/10.1002/cne.920180503

1411