



Sensory attenuation is modulated by the contrasting effects of predictability and control

Anthony W. Harrison^{a,*}, Damien J. Mannion^a, Bradley N. Jack^b, Oren Griffiths^c, Gethin Hughes^d, Thomas J. Whitford^a

^a School of Psychology, University of New South Wales, Sydney, Australia

^b Research School of Psychology, The Australia National University, Canberra, Australia

^c College of Education, Psychology and Social Work, Flinders University, Adelaide, Australia

^d Department of Psychology, University of Essex, Colchester, United Kingdom

ARTICLE INFO

Keywords:

Sensory attenuation
Temporal predictability
Temporal control
Auditory Evoked Potential

ABSTRACT

Self-generated stimuli have been found to elicit a reduced sensory response compared with externally-generated stimuli. However, much of the literature has not adequately controlled for differences in the temporal predictability and temporal control of stimuli. In two experiments, we compared the N1 (and P2) components of the auditory-evoked potential to self- and externally-generated tones that differed with respect to these two factors. In Experiment 1 ($n = 42$), we found that increasing temporal predictability reduced N1 amplitude in a manner that may often account for the observed reduction in sensory response to self-generated sounds. We also observed that reducing temporal control over the tones resulted in a reduction in N1 amplitude. The contrasting effects of temporal predictability and temporal control on N1 amplitude meant that sensory attenuation prevailed when controlling for each. Experiment 2 ($n = 38$) explored the potential effect of selective attention on the results of Experiment 1 by modifying task requirements such that similar levels of attention were allocated to the visual stimuli across conditions. The results of Experiment 2 replicated those of Experiment 1, and suggested that the observed effects of temporal control and sensory attenuation were not driven by differences in attention. Given that self- and externally-generated sensations commonly differ with respect to both temporal predictability and temporal control, findings of the present study may necessitate a re-evaluation of the experimental paradigms used to study sensory attenuation.

1. Introduction

The term ‘sensory attenuation’ has been used to describe a reduction in the neurophysiological response and subjective intensity of sensations elicited by self-generated stimuli, as compared to that of physically identical stimuli when externally-generated (e.g., Blakemore et al., 1998; Schafer and Marcus, 1973). This phenomenon has been argued to involve the operation of internal forward models (Miall and Wolpert, 1996), in which duplicates of the motor commands (‘efference copies’) are used to predict and suppress the sensory consequences of action (Ford and Mathalon, 2012). In this way, sensory attenuation is believed to occur when one’s predictions regarding the outcome of an action match the actual sensory feedback (Wolpert, 1997). Research has explored the effects of sensory attenuation across a variety of action types and sensory modalities, including visual continuity during saccadic eye movements (Bridgeman, 1995; Thakkar et al., 2015), suppression of self-generated speech (Houde et al., 2002; Whitford, 2019)

and the inability to tickle oneself (Blakemore et al., 1998). Research has also observed sensory attenuation based on less innate action-stimulus pairs, including computer-generated tones triggered by a button press (Aliu et al., 2009; Klaffehn et al., 2019; Lange, 2011; Martikainen et al., 2005; Sato, 2008; Sowman et al., 2012). However, despite an extensive corpus of research investigating its mechanisms (see Horváth, 2015), there are at least two potential methodological confounds that draw into question the existence and magnitude of the sensory attenuation effect (Hughes et al., 2013). In particular, much of the sensory attenuation literature has inadequately controlled for the effects of temporal predictability and temporal control.

Temporal predictability – the ability to predict the onset of a stimulus – is intrinsic to many self-generated sensations, which often arise more-or-less instantaneously from one’s own motor actions. Externally-generated stimuli may also be made temporally predictable by providing cues to indicate their impending onset. Research has demonstrated that increasing the temporal predictability of an externally-generated stimulus reduces both its subjective intensity and neurophysiological re-

* Corresponding author.

E-mail address: anthony.harrison@unsw.edu.au (A.W. Harrison).

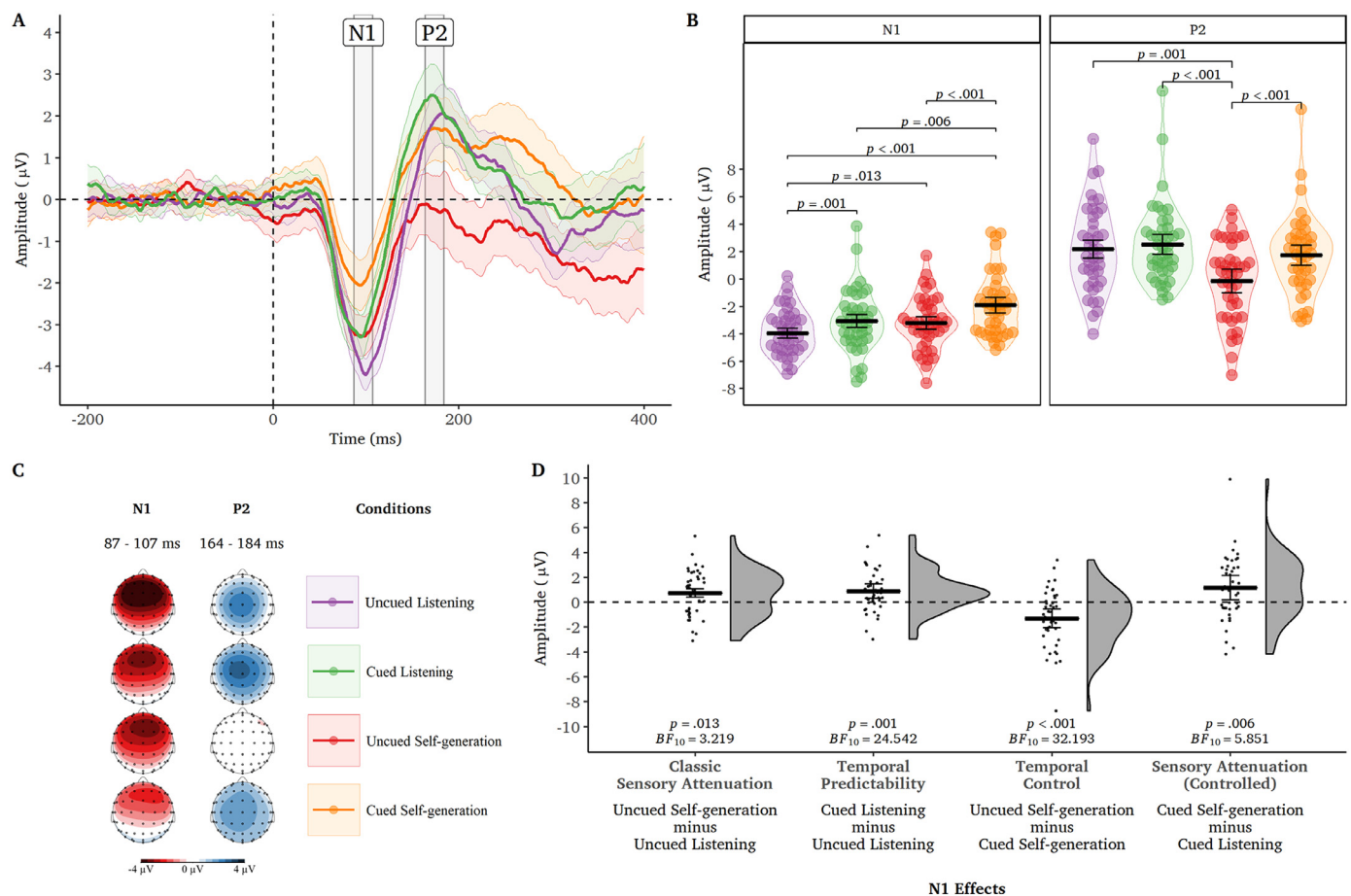


Fig. 2. Results from Experiment 1.

A. Auditory evoked potentials for Experiment 1, including pooled mean amplitudes at Fz, FCz and Cz, as well as ribbons representing 95% CIs. **B.** Mean voltages and 95% CIs for N1 (left) and P2 (right) by condition. Significant contrasts are indicated with their corresponding p -value. N1 amplitudes reflect pooled recordings at Fz, FCz and Cz, while P2 amplitudes reflect those at FCz, Cz and CPz. **C.** Topographic voltage maps for N1 (left) and P2 (right) components, with corresponding condition labels and legend for panels A and B. **D.** Key within-subject contrasts of N1 amplitude with mean difference and 95% CIs. These include contrasts illustrating the sensory attenuation effect (i.e., self- vs. external) without controlling for temporal predictability or control (*uncued self-generation vs uncued listening*), the effect of temporal predictability (*cued listening vs uncued listening*), the effect of temporal control (*uncued self-generation vs cued self-generation*) and sensory attenuation controlling for both temporal predictability and control (*cued self-generation vs cued listening*).

2.1. Method

2.1.1. Participants

The final sample consisted of 42 healthy participants (25 females), aged between 17 and 33 years ($M = 20.69$, $Mdn = 19.24$, $SD = 3.71$). Data from an additional participant was found to have an insufficient number of valid segments following artefact rejection, based on a minimum threshold of 30 usable trials per condition (see *EEG Processing and Analysis*), and was excluded from further analysis. Participants were recruited through the University of New South Wales (UNSW) electronic participants recruitment system (SONA-1) and received course credit in exchange for their participation. The study was approved by the UNSW Human Research Ethics Advisory Panel (Psychology).

2.1.2. Materials and design

Participants were seated facing a BenQ XL2420T computer monitor (24-inch, 1920×1080 resolution screen) at a distance of approximately 60 cm. They were fitted with Sennheiser HD201 headphones and an EEG recording cap, containing a BioSemi ActiveTwo system on which 64 Ag/AgCl active electrodes were positioned according to the extended 10–20 system. Recordings from an electrode placed below the left eye was used in conjunction with Fp1 to develop a vertical electro-oculogram (EOG), and a horizontal EOG was constructed using elec-

trodes placed adjacent to the outer canthus of each eye. Electrodes were also placed on the left and right mastoids, as well as the tip of the nose. The CMS and DRL electrode sites were used as ground electrodes during data acquisition, which was conducted at a sampling rate of 2048 Hz.

In each experimental block, participants watched a visual animation. Against a black background, a red (fixation) line was positioned at the centre of the screen. This vertical line was one pixel wide and approximately six centimetres in length (i.e., 5.4 degrees of visual angle). Participants were instructed to fix their gaze on the centre of this line for each of the eight experimental conditions. In four *uncued* conditions (i.e., *active uncued*, *motor uncued* and *passive* and *visual*), a randomly spaced sequence of white line fragments appeared from offscreen on the right of the monitor and moved leftward across the screen at a constant velocity of $3^\circ/\text{s}$ (see Fig. 1B). These fragments were equal in length and distributed in five adjacent rows spanning the height of the fixation line. After crossing behind the fixation line, the white line fragments continued leftward off the screen. In the *active uncued* and *motor uncued* conditions, participants were asked to repeatedly press a keyboard button, at will. They were asked to aim for an interval of approximately two to four seconds between each button-press and to make their timing “as random and unpredictable as possible”. In the *active uncued* condition, each button press elicited an 85 dB SPL (A-weighted) pure tone (1000 Hz, 10 ms ramp, 100 ms duration), delivered via headphones. For low-latency de-

Table 1
Analysis condition characteristics and their contrasts.

	Self-generated	Temporal predictability	Temporal control	Contrasts		
				Uncued listening	Cued listening	Uncued self-generation
Uncued listening						
Cued listening		✓		TP		
Uncued self-generation	✓	✓	✓	SG, TP, TC	SG, TC	
Cued self-generation	✓	✓		SG, TP	SG	(TC)

Note. Contrasts represent the characteristics of row conditions relative to those in columns. SG= self-generation, TP = temporal predictability and TC = temporal control. Brackets denote relative absence of characteristic in row condition.

livery of auditory stimuli, these were delivered via an AudioFile Stimulus Processor (Cambridge Research Systems). In the *passive uncued* and *visual uncued* conditions, participants were not instructed to press any keys but to remain focussed on the fixation line while stimuli were presented to them. Identical tones were delivered in the *passive uncued* condition based on inter-stimulus intervals (ISIs) established during the *active uncued* condition, which was always presented first in each block. Participants were presented with silent audio tracks in the *visual uncued* and *motor uncued* conditions, with the ISIs based on the *active uncued* condition.

In the four *cued* conditions (i.e., *action cued*, *motor cued*, *passive cued* and *visual cued*), vertical white lines with equal length and width to the fixation line appeared from offscreen to the right of the monitor and moved leftward across the screen at a constant rate (see Fig. 1B). These lines were also spaced according to the timing established by participants through their button presses in the *active uncued* condition. In the *active cued* and *motor cued* conditions participants were instructed to press a key at the precise moment that each stimulus line intersected with the fixation line. A tone was delivered each time the key was pressed in the *active cued* condition, while silent audio tracks were presented in the *motor cued* condition. Participants were not asked to press the keyboard button in the *passive cued* or *visual cued* conditions. In the *passive cued* condition, a tone was presented each time a vertical white line passed the fixation line, while participants were presented with silent audio tracks in the *visual cued* condition. Audio tracks for the *passive* and *visual* conditions were also delivered via the AudioFile Stimulus Processor.

Trials were divided into sets of eight blocks, with 30 trials per block. Participants underwent three sets, meaning that there were 90 trials of each of the eight conditions in total. Each set commenced with an *active uncued* block, followed by one block of each other condition presented in random order. Data collection lasted approximately 50 min, and included short breaks between blocks.

The auditory evoked potentials for each condition were corrected by subtracting the ERPs for tasks that did not involve auditory stimuli but were otherwise identical. These *visual* and *motor* conditions were used to correct auditory evoked potentials of the *passive* and *active* conditions, respectively. There were thus eight block types in total: *uncued* and *cued* condition variants of *passive*, *active*, *visual*, and *motor*. These eight experimental conditions supported development of four analysis conditions: *uncued listening*, *cued listening*, *uncued self-generation* and *cued self-generation*, where *listening* conditions represented visual-corrected *passive* conditions and *self-generation* conditions represented motor-corrected *active* conditions (see Fig. 1C).

The corrected waveforms allow comparison of self- and externally-generated tones in a manner that accounted for both temporal predictability and temporal control (see Table 1). Four contrasts hold particular relevance. Firstly, comparison of *uncued listening* and *cued listening* allowed evaluation of the effects of *temporal predictability*. Secondly, the effect of *temporal control* could be assessed through comparison of *cued self-generation* and *uncued self-generation*. Thirdly, the *classic sensory attenuation* effect (i.e., without controlling for the effects of temporal predictability or temporal control) was demonstrated through comparison of *uncued listening* and *uncued self-generation*. Finally, the *sensory*

attenuation (controlled) effect (i.e., controlling for both temporal predictability and temporal control) was evaluated based on comparison of *cued listening* and *cued self-generation*.

2.1.3. EEG processing and analysis

Data were referenced to the nose electrode and filtered offline using BrainVision Analyzer. A notch filter (50 Hz) was applied, as well as a phase-shift free half-amplitude Butterworth band-pass filter (0.1 Hz to 30 Hz) with 48 dB/Oct slope. Data were then segmented into 600 ms epochs, which included 200 ms prior to sound stimulus presentation and 400 ms post-onset. Eye movement artefacts were corrected using the method described by Miller et al. (1988), based on the approach developed by Gratton et al. (1983). Segments found to contain peak-to-peak amplitudes in excess of 200 μ V were excluded. Baseline correction was applied using the average voltage in the 200 ms prior to stimulus onset.

The primary dependent variable was the amplitude of the auditory N1 component, while the P2 was a secondary component of interest. The N1 component of the auditory event-related potentials (ERPs) was analysed using pooled recordings taken at electrode sites Fz, FCz, and Cz, reflecting its frontocentral topography. The P2 component is known to have a more central distribution, and was therefore analysed using recordings taken at FCz, Cz and CPz. These were locked to auditory stimulus onset in the *passive* and *active* conditions. In the *motor* and *visual* conditions, segments were time-locked to the onset of each silent audio track. These conditions differed from the *active* and *passive* conditions only in that the audio track was silent, thereby supporting correction for the effects of motor action and visual animation respectively. Specifically, the *active* conditions were corrected by subtracting the ERPs of corresponding motor condition (i.e., *cued* and *uncued*), while the *passive* conditions were corrected by subtracting the ERPs of the corresponding visual condition.

Grand average waveforms were calculated based on at least 30 usable trials for each of the eight conditions. A one-way repeated measures ANOVA was used to assess potential differences in the number of trials included in grand average waveforms between conditions. This included trial counts for *active uncued* ($M = 87.64$, $SD = 3.46$), *active cued* ($M = 87.62$, $SD = 4.10$), *motor uncued* ($M = 87.31$, $SD = 4.64$), *motor cued* ($M = 85.55$, $SD = 9.37$), *passive uncued* ($M = 86.40$, $SD = 6.43$), *passive cued* ($M = 86.10$, $SD = 8.28$), *visual uncued* ($M = 84.12$, $SD = 11.62$) and *visual cued* ($M = 84.60$, $SD = 9.91$). Non-sphericity was identified by Mauchly's test (Mauchly, 1940) and corrected using the Greenhouse-Geisser method (Geisser and Greenhouse, 1958). Results did not reveal a statistically significant difference in the number of trials between conditions, $F(3.63, 148.72) = 2.196$, $p = 0.079$, partial $\eta^2 = 0.051$, $BF_{10} = 0.45$.

The N1 is known to have a frontocentral maximum (Zouridakis et al., 1998), which was consistent with observations in the present data (maximal at FCz). A more central P2 (maximal at Cz) was also consistent with expectations based on previous research (Potts et al., 1998). Analyses of the N1 component were based on pooled recordings at electrode sites Fz, FCz and Cz, as has been done previously in recognition of the maximal N1 signal at these locations (Jack et al., 2019; Näätänen and Picton, 1987; Whitford et al., 2017; Woods, 1995). Similarly, P2 compo-

nent amplitudes were based on pooled recording at electrode sites FCz, Cz and CPz. Both N1 and P2 components were identified based on a collapsed localizer waveform (i.e., averaging across all participants and conditions; Luck and Gaspelin, 2017). The N1 component was identified as the most negative local minimum between 25 ms and 175 ms in the collapsed localizer waveform, which was identified at 97 ms post-stimulus. Mean amplitudes were calculated for each condition based on 20 ms windows centred on this peak (i.e., between 87 and 107 ms post-stimulus). The P2 component was identified as the most positive local maximum between 110 ms and 200 ms, which was found to occur at 175 ms in the collapsed localizer waveform. The mean amplitude of P2 components were consequently calculated based on sampled voltage between 165 ms and 185 ms.

2.1.4. Statistical analysis

One-way repeated measures analyses of variance (ANOVAs) were undertaken to assess the statistical significance of differences between component voltages observed in each condition. This statistical approach was selected on the basis that although the experimental design involved two factors (i.e., active/passive and cued/uncued), change between levels in one factor did not equate to the same conceptual change between levels in the other. For example, cues in the *passive* and *visual* conditions supported temporal predictability of stimuli while those in the *active* and *motor* conditions affected only temporal control (see Table 1). A 2×2 factorial ANOVA would therefore have risked conflating these effects. The Greenhouse-Geisser method (Geisser and Greenhouse, 1958) was used to correct for non-sphericity where this was identified by Mauchly's tests (Mauchly, 1940).

A series of paired samples Student's *t*-tests were conducted to evaluate differences between the mean voltage observed for components of each condition. The Benjamini-Hochberg method (Benjamini and Hochberg, 1995) was used to correct *p*-values so as to control the False Discovery Rate (FDR). Contrasts were only interpreted (i.e., using uncorrected *p*-values) when these remained significant following correction for multiple comparisons (i.e., $p_{corr} < 0.05$; see Appendix Table 1). Where correction resulted in a contrast falling outside the critical *p*-value, this was explicitly acknowledged. Corrections were applied to control FDR by experiment (i.e., 12 contrasts in Experiment 1 and 30 contrasts in Experiment 2, see below). A Bayes factor (BF_{10}) was also developed for omnibus repeated measures ANOVAs, as well as each contrast. These comparisons included Cauchy priors with an *r*-scale of $1/\sqrt{2}$ for effect size (see Morey and Rouder, 2015), which have been recommended across a range of Bayesian statistical procedures, including regression (Gelman et al., 2008) and point null hypothesis testing (Jeffreys, 1998).

The accuracy of button presses in the *active cued* and *motor cued* conditions were compared using a paired samples Student's *t*-test. Inter-stimulus intervals (ISIs) were also analysed to confirm effective and consistent manipulation of timing in conditions involving participant input. Paired samples Student's *t*-tests were used to compare mean ISIs in *active uncued* and *active cued*, as well as the motor-only conditions used for their correction (i.e., *motor uncued* and *motor cued*). Results from analyses of ISIs are provided in *Supplementary Online Material*.

2.2. Results

2.2.1. N1

A one-way repeated measures ANOVA was conducted to assess differences in N1 amplitude (i.e., between 87 and 107 ms post-stimulus). Results revealed a statistically significant difference between conditions, $F(2.51, 103.01) = 13.171$, $p < 0.001$, partial $\eta^2 = 0.243$, $BF_{10} = 91,358.45$. A series of paired samples Student's *t*-tests were conducted to compare participants' mean N1 amplitudes across conditions. Results revealed that N1 amplitudes (μV) in the *uncued listening* condition ($M = -3.96$, $SD = 1.78$) were significantly more negative than those in each of the other conditions, including *cued listening* ($M = -3.07$,

$SD = 2.24$), $t(41) = -3.46$, $p = 0.001$, $d = 0.44$, $BF_{10} = 24.54$, *uncued self-generation* ($M = -3.22$, $SD = 1.93$), $t(41) = -2.60$, $p = 0.013$, $d = 0.40$, $BF_{10} = 3.22$, and *cued self-generation* ($M = -1.91$, $SD = 2.26$), $t(41) = -6.10$, $p < 0.001$, $d = 1.01$, $BF_{10} = 50,782.42$ (see Appendix Table 1 for detail). N1 amplitudes of the *cued self-generation* condition were also found to be less negative than those of both *cued listening*, $t(41) = 2.87$, $p = 0.006$, $d = 0.51$, $BF_{10} = 5.85$, and *uncued self-generation*, $t(41) = 3.57$, $p < 0.001$, $d = 0.62$, $BF_{10} = 32.19$. Amplitudes of the N1 in the *uncued self-generation* and *cued listening* conditions were not found to differ significantly, $t(41) = 0.47$, $p = 0.639$, $BF_{10} = 0.19$.

2.2.2. P2

A one-way repeated measures ANOVA identified significant differences between the mean amplitudes (μV) of P2 components across conditions, $F(3, 123) = 9.486$, $p < 0.001$, partial $\eta^2 = 0.188$, $BF_{10} = 4902.41$. Pairwise comparisons suggested that the amplitudes of P2 in the *uncued self-generation* condition ($M = -0.14$, $SD = 2.88$) were significantly less than those of the *cued self-generation* ($M = 1.73$, $SD = 2.96$), $t(41) = -3.46$, $p = 0.001$, $d = 0.64$, $BF_{10} = 24.31$, *uncued listening* ($M = 2.18$, $SD = 3.10$), $t(41) = -4.44$, $p < 0.001$, $d = 0.78$, $BF_{10} = 352.03$, and *cued listening* ($M = 2.52$, $SD = 2.94$), $t(41) = -4.34$, $p < 0.001$, $d = 0.92$, $BF_{10} = 261.99$, conditions. The P2 amplitudes of the other conditions did not significantly differ from each other (see Appendix Table 2 for detail).

2.2.3. Behavioural data

A paired samples Student's *t*-test was used to compare participants' accuracy in pressing the keyboard button to synchronise with visual stimuli in the two cued conditions. Results suggested that mean error (*ms*; i.e., averaged across all trials for each participant) in the *active cued* condition ($M = -73.36$, $SD = 41.15$, max = 104.39) and the *motor cued* condition ($M = -81.47$, $SD = 40.72$, max = 96.24) did not differ significantly from each other, $t(41) = 0.93$, $p = 0.358$, $BF_{10} = 0.25$.

2.2.4. Power analyses

Post hoc analyses explored the power of the sample in Experiment 1 ($n = 42$) to detect small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$) effects, according to standardised reporting conventions (Cohen, 1992). Power (β) varied based on the impact of correction for multiple comparisons, and included a range for small (0.05 to 0.24), medium (0.58 to 0.89) and large (0.98 to 1.00) effects.

2.3. Discussion

Through analysis of the contrasts, we were able to isolate the effects of temporal predictability and temporal control from one's input to generating sensations (see Table 1). As hypothesised, N1 amplitude in the *cued listening* condition was found to be significantly less negative than in *uncued listening*. This is consistent with previous research demonstrating that an increase in the temporal predictability of stimulus is associated with a reduction in neurophysiological response (e.g., Lange, 2009; Schafer and Marcus, 1973; Weiskrantz et al., 1971). In addition, the classic sensory attenuation effect was replicated in that smaller N1 amplitudes were observed in the *uncued self-generation* condition compared with *uncued listening*. Despite this, the N1 amplitude of the *uncued self-generation* condition was not found to differ significantly from that of *cued listening*, and the associated Bayes factor provided substantial evidence in favour of the null hypothesis (Jeffreys, 1998). This result suggests that the classic sensory attenuation effect may be eliminated by controlling for temporal predictability without also accounting for temporal control.

Critically, reduced N1 amplitudes were observed in the *cued self-generation* condition compared with *cued listening*. A reduced neurophysiological response to self-generated stimuli was therefore evident when the *self-generation* and *listening* conditions were matched in terms of both

temporal predictability and temporal control. This result suggests that the phenomenon of sensory attenuation prevails when controlling for differences in both temporal predictability and temporal control, but may otherwise be conflated with the effect of temporal predictability alone. Of significance, N1 amplitudes elicited by tones in the *cued self-generation* condition were significantly smaller than in the *uncued self-generation* condition, owing to an apparent amplification related to temporal control. As reflected in the contrast between *uncued self-generation* and *cued listening*, the scale of this amplification was such that sensory attenuation (i.e., the difference in neurophysiological response between self- and externally generated stimuli) was diminished for stimuli over which participants exerted temporal control.

Although the functional significance of the P2 is relatively poorly understood, research has identified that it is influenced by factors that are both common to and differentiable from those affecting the N1. Heightened attention to stimuli has long been associated with a negativity that manifests as simultaneous enhancement of the N1 component and suppression of the P2; an effect described variously as ‘Processing Negativity’ (Näätänen et al., 1978) and ‘Nd’ (Hillyard et al., 1973). More negative N1 and P2 amplitudes in the *uncued self-generation* condition, relative to *cued self-generation*, may thereby reflect heightened auditory attention when participants had control over the timing of the stimuli. Although recent evidence suggests that P2 amplitude may be positively correlated with one’s sense of agency over sound (Timm et al., 2016), it is noteworthy that control over the timing of stimuli was associated with smaller P2 amplitudes in the present study. A potential reason for this discrepancy is that the study by Timm et al. (2016) used an illusion to examine the effects of perceived ownership (i.e., by making self-generated tones appear as if they were not a result of participants’ actions). In contrast, the present investigation explored effects involving temporal control (i.e., agency in determining when to generate a stimulus; see Haggard, 2017). Significantly, Timm et al. (2016) also observed reduced P2 amplitudes for self-generated tones compared with externally-generated tones when participants were not subject to the illusion of non-ownership.

Potential discrepancies in the allocation of selective attention present a possible confound to the results of Experiment 1. Previous research has demonstrated amplification of the auditory N1 in tasks that required attention to auditory stimuli, compared to conditions in which attention was directed to visual stimuli (Hackley et al., 1990; Woods et al., 1992). If the *cued self-generation* condition involved heightened attention to visual stimuli in order to accurately time each button press, compared with *cued listening*, this may account for a reduced N1 amplitude in the former. To assess this possibility, the second experiment sought to both replicate key findings of Experiment 1 and quantify the effect of manipulating selective attention to the visual stimuli.

3. Experiment 2

In Experiment 2, we aimed to investigate the effect of selective attention on the auditory N1 amplitude. We focussed, in particular, on the effects of selective attention in differences between the *cued listening* and *cued self-generation* conditions. This contrast, which we describe as *sensory attenuation (controlled)*, allows direct comparison of sensory responses to self- and externally-generated stimuli while holding constant the level of temporal predictability and temporal control. In Experiment 2, the potential effect of selective attention on the auditory evoked potential was explored by modifying task requirements such that similar levels were allocated to visual stimuli in variants of both the *passive cued* and *active cued* experimental conditions.

A proportion of the lines/line segments in the visual stimuli were changed from white to light grey (see Fig. 3B). With this exception, the *active uncued*, *active cued* and *passive cued* conditions were collected with identical procedures to Experiment 1. An additional two conditions were included that required participants to keep a tally of the number of

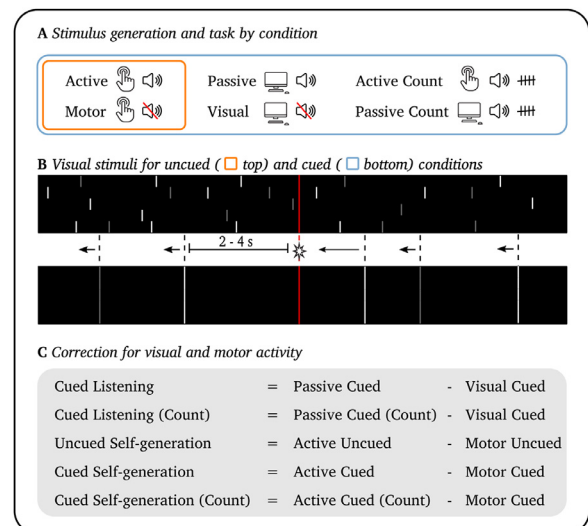


Fig. 3. Experiment 2 visual stimuli and schematic.

A. Participants pressed a keyboard button to initiate events in the *active*, *active (count)* and *motor* conditions, while events were externally-generated (by computer) in the *passive*, *passive (count)* and *visual* conditions. Events involved an 85 dB tone (1000 Hz) in the *active*, *active (count)*, *passive* and *passive (count)* conditions, while silent audio tracks marked events in the *motor* and *visual* conditions. Orange and blue boxes indicate the stimulus generation conditions (panel A) that were paired with *uncued* and *cued* conditions, respectively (panel B). *Uncued* and *cued* variants were presented for *active* and *motor* conditions, while other conditions involved only *cued* variants. **B.** The two *uncued* conditions were identical to corresponding conditions in Experiment 1, except that half of the line fragments were grey while the other half were white. These were the *active uncued* and *motor uncued* conditions. In six *cued* conditions, a proportion of lines were also made grey. This was done such that half of the lines were grey across the course of the experiment, while the exact proportion varied slightly within individual blocks. The *active cued*, *motor cued*, *passive cued* and *visual cued* conditions were otherwise identical to Experiment 1, with participants instructed to ignore variation in line shading. The *active cued (count)* and *cued passive (count)* conditions were similar to *active cued* and *passive cued*, respectively, except that participants were required to keep a mental tally of the number of target lines (i.e., white or grey) and report this at the conclusion of the block. **C.** The eight experimental conditions supported development of five analysis conditions: *cued listening*, *cued listening (count)*, *uncued self-generation*, *cued self-generation*, and *cued self-generation (count)*.

long white or grey lines, and to report this at the conclusion of relevant blocks. These included the *active cued (count)* and *passive cued (count)* conditions. The contrast of the grey lines was such that they were difficult to distinguish until they approached the fixation line.

An analysis of the anterior N2 component (N2b) was used as a manipulation check of increased attentional load in the counting task. This component is observed in frontocentral regions between 200 ms and 350 ms post-stimulus on tasks that involve cognitive control and manipulations of cognitive load (Folstein and Van Petten, 2008). The N2b can be distinguished from other components by its sensitivity to the level of attention allocated to stimuli (Pritchard et al., 1991) and insensitivity to stimulus probability (Luck and Hillyard, 1994). In addition, the N2b has been associated with response inhibition tasks, including the ‘go/no-go’ paradigm (Donders, 1969), in which participants are required to respond to one stimulus while withholding their response to others (Bruin and Wijers, 2002). Importantly, it has also been found to be influenced by similar tasks requiring only mental responses, such as counting target stimuli (Mertens and Polich, 1997). We therefore hypothesised that conditions requiring participants to keep tally of the number of target lines would elicit larger N2b components than conditions that did not involve counting.

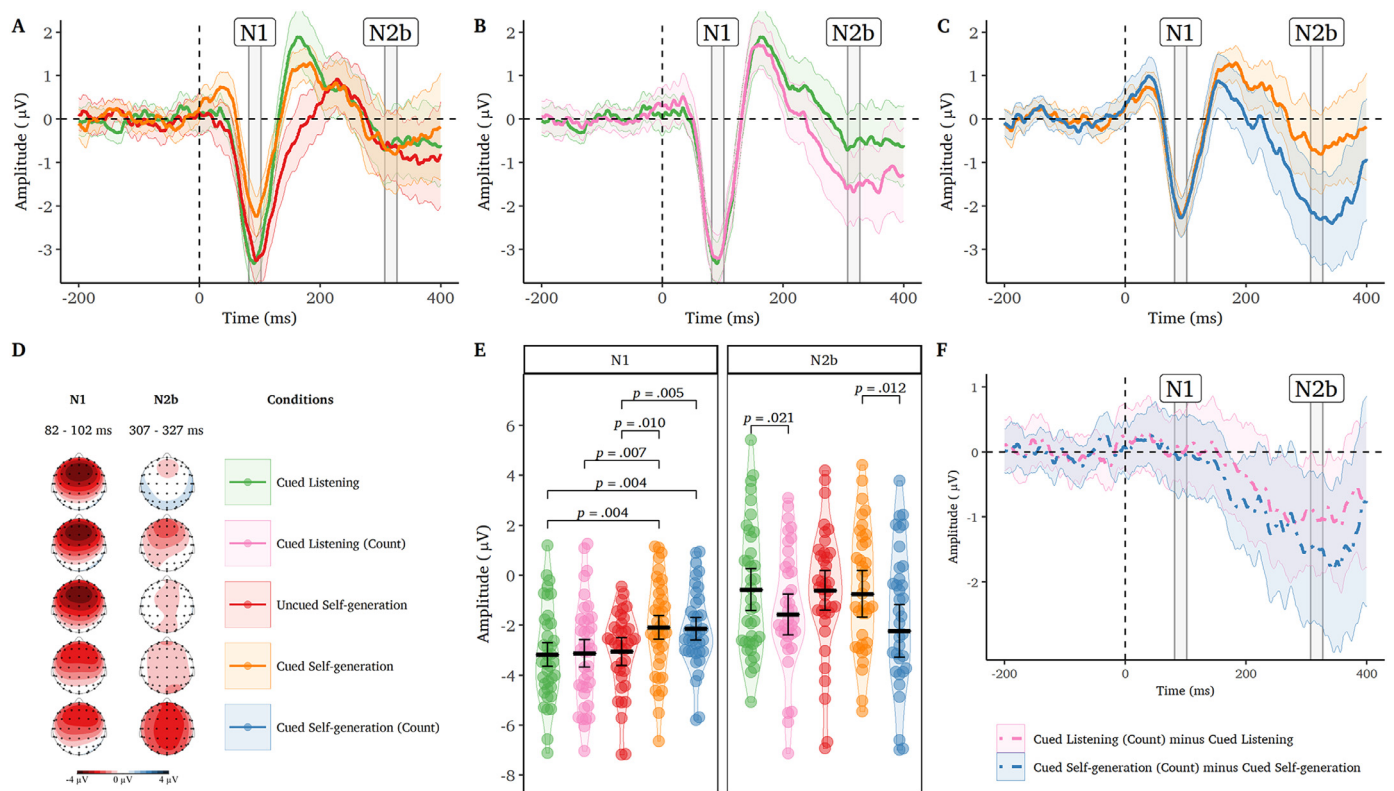


Fig. 4. Results from Experiment 2.

A to C. Auditory evoked potentials for Experiment 2, including pooled mean amplitudes of Fz, FCz and Cz, as well as ribbons representing 95% CIs. Panels present (A) conditions replicating those of Experiment 1, including *cued listening*, *uncued self-generation* and *cued self-generation*, (B) *cued listening* and *cued listening (count)* (C) *cued self-generation* and *cued self-generation (count)*. **D.** Topographic voltage maps for the N1 and N2b latency window, with corresponding condition labels and legend for panels A to C and E to F. Significant contrasts are indicated with their corresponding p -value. **E.** Mean voltages and 95% CIs for the N1 (left) and N2b (right) components, by condition. **F.** Difference waves comparing *cued listening (count)* and *cued self-generation (count)* with their non-counting equivalents (*cued listening* and *cued self-generation*, respectively).

3.1. Method

3.1.1. Participants

A further 38 healthy participants (25 females), aged between 17 and 36 years ($M = 21.38$, $Mdn = 19.22$, $SD = 4.80$), were recruited for Experiment 2. Participants were again recruited through the University of New South Wales (UNSW) electronic participants recruitment system (SONA), with approval provided for the study by the UNSW Human Research Ethics Advisory Panel (Psychology). Data from an additional four participants were collected but excluded from further analysis due to self-reported recreational drug use within the preceding 48 h (three participants) and self-reported diagnosis of a psychotic disorder (one participant).

3.1.2. Materials and design

EEG recording, data processing and statistical analyses were undertaken according to the same specifications as described for Experiment 1. Visual stimuli were adjusted, such that a proportion of white lines and line fragments were replaced by grey lines (see Fig. 3B). As with Experiment 1, a total of 90 trials were presented for each condition. Across three blocks, each containing 30 trials, a total of 14, 15 and 16 white lines (or equivalent line fragments) were presented. The *active uncued*, *active cued* and *passive cued* conditions, as well as their *motor* and *visual* controls, were collected according to the procedures described for Experiment 1. In these conditions, participants were instructed to disregard variation in line colour. Two additional conditions were included, which asked participants to count the total number of lines of a particular variety (i.e., white or grey) in addition to existing requirements. Shading of the white and grey lines was made such that they were diffi-

cult to distinguish until they approached the fixation line, and the order of white and grey lines was randomised within blocks.

Blocks were presented in three sets, each containing one block for each of the eight experimental conditions. The order of blocks was again pseudorandom, such that each set commenced with the *active uncued* condition. The line shade ratio of blocks was randomised across sets (i.e., different conditions were able to contain blocks with differing ratios for each set). The target line type (i.e., white or grey) was counterbalanced across participants. As such, eight separate conditions were administered, including two *uncued* (i.e., *active* and *motor*) and six *cued* (i.e., *active*, *passive*, *motor*, *visual*, *active count* and *passive count*). The waveforms for *active* and *passive* conditions (i.e., including the *count* conditions) were corrected for *motor* and *visual* effects, respectively, using the same method as described for Experiment 1. The resulting analysis conditions included *self-generation uncued*, *listening cued*, *listening cued (count)*, *self-generation cued*, and *self-generation cued (count)*.

As in Experiment 1, at least 30 trials were obtained per participant for each of the eight conditions. A one-way repeated measures ANOVA was used to assess potential differences in the number of trials included in grand average waveforms between conditions. This included trial counts for *active uncued* ($M = 88.53$, $SD = 4.05$), *active cued* ($M = 89.05$, $SD = 2.86$), *active cued count* ($M = 88.11$, $SD = 5.08$), *motor uncued* ($M = 87.03$, $SD = 6.92$), *motor cued* ($M = 89.34$, $SD = 2.22$), *passive cued* ($M = 89.00$, $SD = 2.25$), *passive cued count* ($M = 89.32$, $SD = 1.36$) and *visual cued* ($M = 88.55$, $SD = 2.72$). Non-sphericity was identified by Mauchly's test (Mauchly, 1940) and corrected using the Greenhouse-Geisser method (Geisser and Greenhouse, 1958). Results did not reveal a

statistically significant difference in the number of trials between conditions, $F(3.23, 119.48) = 1.98, p = 0.117$, partial $\eta^2 = 0.051$, $BF_{10} = 0.31$.

The N1 component was identified as the most negative local minimum between 25 ms and 175 ms in the collapsed localizer waveform. This was found to occur 92 ms post-stimulus, with mean amplitudes calculated between 82 ms and 102 ms. The P2 component was again identified in the collapsed localizer waveform as the most positive local maximum between 110 ms and 200 ms, which was found to occur at 177 ms. P2 component amplitudes were therefore calculated based on recordings between 167 ms and 187 ms. Although P2 analyses are not reported in the main body, descriptive statistics and contrasts of each condition may be found in [Appendix Table 4](#). The N2b component was analysed using pooled recordings at the Fz, FCz and Cz electrode sites, reflecting its frontocentral topography. These sites were selected on the basis that, while varied, investigations of the anterior N2 have typically utilised frontal locations on the midline of the scalp ([Näätänen and Picton, 1986](#)). The N2b components were defined based on the collapsed localizer method, using pooled recordings from electrode sites Fz, FCz and Cz between 200 ms and 350 ms post-stimulus. The N2b component was found to occur at 317 ms, supporting calculation of component amplitudes between 307 ms and 327 ms.

3.2. Results

3.2.1. N1

The N1 peak was identified as having occurred 92 ms post-stimulus. Amplitudes of the N1 component (i.e., between 82 and 102 ms) were compared using a one-way repeated measures ANOVA. Results revealed a statistically significant difference between conditions, $F(3.11, 115.16) = 4.944, p = 0.002$, partial $\eta^2 = 0.118$, $BF_{10} = 30.23$. A series of paired samples Student's *t*-tests identified that the mean amplitude (μ V) of the *uncued self-generation* condition ($M = -3.05, SD = 1.61$) was significantly more negative than those of the *cued self-generation* condition ($M = -2.09, SD = 1.89$), $t(37) = -2.70, p = 0.010, d = 0.55, BF_{10} = 4.06$, and the *cued self-generation (count)* condition ($M = -2.14, SD = 1.57$), $t(37) = -3.00, p = 0.005, d = 0.57, BF_{10} = 7.87$ (see [Appendix Table 3](#) for detail). Similarly, the *cued listening* condition ($M = -3.18, SD = 1.88$) was found to have larger (i.e., more negative) N1 components than those of the *cued self-generation* condition, $t(37) = -3.08, p = 0.004, d = 0.58, BF_{10} = 9.34$, and *cued self-generation (count)* condition, $t(37) = -3.09, p = 0.004, d = 0.60, BF_{10} = 9.50$. The *cued listening (count)* condition ($M = -3.12, SD = 2.03$) was also found to have larger N1 components than those of the *cued self-generation* condition, $t(37) = -2.85, p = 0.007, d = 0.52, BF_{10} = 5.55$. However, the difference between N1 amplitudes in the *cued listening (count)* condition was not found to differ significantly from the *cued self-generation (count)* condition after correcting for multiple comparisons, $t(37) = -2.47, p = 0.018, p_{corr} = 0.055, d = 0.54, BF_{10} = 2.50$. The N1 amplitudes of *cued self-generation* and *cued self-generation (count)* were not found to differ significantly, $t(37) = 0.18, p = 0.856, BF_{10} = 0.18$, nor were those of *cued listening* and *cued listening (count)*, $t(37) = -0.22, p = 0.828, BF_{10} = 0.18$. Bayes factors representing the comparison of *cued self-generation* and *cued listening* conditions with their counting equivalents both represent substantial evidence in favour of null hypotheses ([Jeffreys, 1998](#)).

3.2.2. Anterior N2

The anterior N2 component (N2b) peak was found to have occurred 317 ms post-stimulus. Mean amplitudes were calculated in the same manner as for other components, based on a 20 ms window centred on this peak (i.e., between 307 ms and 327 ms). A one-way repeated measures ANOVA identified that the N2b component varied significantly across conditions, $F(2.85, 105.28) = 2.775, p = 0.043$, partial $\eta^2 = 0.07$, $BF_{10} = 1.41$. Pairwise comparison revealed that the mean N2b amplitude (μ V) of the *cued listening (count)* condition ($M = -1.57, SD = 2.81$)

was significantly more negative than those of the *cued listening* condition ($M = -0.57, SD = 2.58$), $t(37) = -2.41, p = 0.021, d = 0.37, BF_{10} = 2.20$ (see [Appendix Table 5](#) for detail). The N2b of the *cued self-generation (count)* condition ($M = -2.23, SD = 3.32$) was found to be significantly more negative than that of the *cued self-generation* condition ($M = -0.75, SD = 2.70$), $t(37) = -2.65, p = 0.012, d = 0.49, BF_{10} = 3.62$. The difference in N2b amplitude between *cued self-generation (count)* and the *uncued self-generation* condition ($M = -0.61, SD = 2.48$) was not found to be statistically significant following correction for multiple comparisons, $t(37) = -2.46, p = 0.019, p_{corr} = 0.051, d = 0.56, BF_{10} = 2.44$. Significant differences were not observed between the *cued listening (count)* condition and the *cued self-generation* condition, $t(37) = -1.20, p = 0.238, BF_{10} = 0.34$, the *cued self-generation (count)* condition, $t(37) = -0.96, p = .826, BF_{10} = 0.27$, or the *uncued self-generation* condition, $t(37) = -1.79, p = 0.082, BF_{10} = 0.21$. In addition, the difference between the *cued listening* condition and the *cued self-generation (count)* condition was not found to be significant following correction for multiple comparisons, $t(37) = 2.13, p = 0.040, p_{corr} = 0.076, BF_{10} = 1.30$.

3.2.3. Behavioural data

Participants' accuracy in keeping tally of the number of target lines in the *active cued (count)* and *passive cued (count)* conditions was calculated as the error rate (i.e., absolute difference between reported and actual number of target lines per block), averaged across the three blocks. Participants were found to demonstrate a high level of accuracy (i.e., low number of errors) in both the *active* ($M = 0.47, SD = 0.55, \max = 2.33$) and *listening* ($M = 0.42, SD = 0.67, \max = 3.33$) variants, and participants' accuracy did not differ significantly between these two conditions, $t(37) = 0.42, p = 0.676, BF_{10} = 0.19$. ISIs and the synchrony of participants' button press to passing lines were assessed in the same manner as Experiment 1 and not found to differ between *cued* conditions (see [Supplementary Online Material](#) for detail).

3.2.4. Power analyses

Post hoc analyses explored the power of the sample in Experiment 2 ($n = 38$) to detect small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$) effects, according to standardised reporting conventions ([Cohen, 1992](#)). Power (β) varied based on the impact of correction for multiple comparisons, and included a range for small (0.02 to 0.22), medium (0.40 to 0.85) and large (0.93 to 1.00) effects.

3.3. Discussion

The effects of *temporal control* and *sensory attenuation (controlled)* were replicated from Experiment 1, with attenuation of N1 amplitudes in the *cued self-generation* condition relative to the *uncued self-generation* and *cued listening* respectively. Differences in N1 amplitude were not observed between the *cued self-generation* and *cued self-generation (count)* conditions, or between the *cued listening* and *cued listening (count)* conditions, and the associated Bayes factors provided substantial evidence in support of the null hypotheses. In contrast, anterior N2 component amplitudes in the *cued self-generation (count)* and *cued listening (count)* conditions were larger (i.e., more negative) than those in equivalent conditions that did not require counting. This was consistent with hypotheses, reflecting effective manipulation of attentional load such that demand for visual attention was greater when participants were required to keep tally of the target lines.

The fact that increased demand for visual attention was not associated with a reduction in N1 amplitude suggests that the observed differences between the *cued listening* and *cued self-generation* conditions are not likely to be driven by differences in selective attention. It is noteworthy that these findings are consistent with previous research that has demonstrated a sensory attenuation effect that is insensitive to manipulation of attention to visual stimuli ([Timm et al., 2013](#)). Insensitiv-

ity of N1 amplitudes to increased visual attention within the *cued self-generation (count)* condition may indicate that the difference between *cued self-generation* and *uncued self-generation* is also unlikely to be attributable to differences in attentional demand. Future research could explore this possibility by using a similar paradigm to draw attention towards visual stimuli during the *uncued self-generation* condition.

4. General discussion

The present study investigated effects associated with temporal predictability and temporal control on the neurophysiological response to self- and externally-generated auditory stimuli. While temporal predictability and temporal control were each shown to affect N1 amplitude, the contrasting nature of these effects meant that the phenomenon of sensory attenuation prevailed when controlling for both. That is, self-generated sounds elicited a smaller cortical response compared with externally-generated sounds when controlling for both temporal predictability and temporal control. Consistent with previous research (Lange, 2009; Schafer and Marcus, 1973; Sowman et al., 2012; Weiskrantz et al., 1971), increasing the temporal predictability of sound stimuli was found to reduce the amplitude of the auditory N1 component. This finding sits within a broader literature demonstrating that the sensory nervous system responds less to stimuli that are more highly predicted (Friston, 2005, 2010). In contrast to the suppressive effect of temporal predictability, a positive relationship was observed between participants' temporal control of self-generated sound and resulting N1 component amplitude. To our knowledge, this effect has not previously been reported. Interestingly, the magnitude of effect was such that the difference in N1 component amplitudes elicited by self- and externally-generated stimuli was diminished when participants had temporal control. We replicated this effect in Experiment 2, which also explored the potential influence of changes in selective attention. We found that changing task requirements to increase selective visual attention in both the *self-generation* and *listening* conditions did not result in a change to the observed pattern of N1 amplitudes.

By suppressing one's neurophysiological response to self-generated stimuli, compared to those created externally, sensory attenuation is believed to serve an adaptive role in removing from our perceptual experience information that can be anticipated on the basis of motor activity alone (Miall and Wolpert, 1996). An observed amplification of temporally-controlled sensations may therefore reflect increased utility of sensory information associated with stimuli that result from deliberately-controlled action, compared with those arising as incidental outcomes of movements. The sounds of one's own footsteps may be of less intrinsic value than someone else's while walking in an alley late at night. However, the sounds generated by one's own footsteps are likely to have more equal importance when attempting to walk quietly across creaky floorboards.

Increased activation of the sensory nervous system during volitional action is also consistent with the tenets of ideomotor theory, which proposes that action is generated through internal activation of its anticipated perceptual consequences (see review by Shin et al., 2010). Within this framework, it has been proposed that action is controlled differently based on whether it is guided by stimulus-response (sensorimotor) learning or through action-effect (ideomotor) learning (Herwig et al., 2007). Our results provide support to this notion, demonstrating increased activation of sensory cortices to stimuli resulting from intention-based action (i.e., *uncued self-generation*) compared to stimulus-based action (i.e., *cued self-generation*). A distinction between these forms of action may also be observed based on the neurological processes involved in their initiation. Recent research has demonstrated increased activation of the corticospinal motor system in response to external cues that have previously been associated with actions elicited by transcranial magnetic stimulation (Tran et al., 2019). Interestingly, this effect has been found

to occur regardless of whether participants expect subsequent action (Tran et al., 2020).

With respect to the neuroanatomy supporting intention-based actions, these appear to be coordinated by structures within the posterior medial frontal cortex (pmFC; see review by Waszak et al., 2012). It is possible that effects associated with temporal control therefore reflect modulation of sensory cortical response by regions within the pmFC, including the supplementary motor area (SMA) and pre-supplementary motor area (pre-SMA). The SMA and pre-SMA have been implicated in a variety of timing functions (Wiener et al., 2010), including the internal production and discrimination of time intervals (Herrmann et al., 2014; Macar et al., 2006). Interestingly, the SMA has been associated with several functions of relevance to the phenomenon of sensory attenuation, including auditory processing (Rauschecker and Scott, 2009) and imagery (McNorgan, 2012), as well as a potential source of motor efference copies (Jürgens, 1984).

A potential alternative interpretation for the effect we have attributed to temporal control is that the effects of prediction based on external cues and self-generation are summative. Such an effect might mean that the difference between *uncued self-generation* and *cued self-generation* is attributable to the additional predictive information provided by external cues, as opposed to differences in temporal control. This may be consistent with recent evidence demonstrating that motor action serves to enhance temporal attention (Zalta et al., 2020) and the accuracy of temporal predictions (Morillon and Baillet, 2017). Future research is therefore needed to explore the manner in which the effects of internal predictions (e.g., based on internal forward models) combine with those of external cues.

Another potential avenue for future research may involve distinguishing effects resulting from volitional motor activation from higher-level appraisal of one's agency in the creation of stimuli. Weiss et al. (2011) proposed that the reduced subjective intensity of self-generated sounds when these were prompted by the experimenter resulted from the social interaction that this involved – an account that favours the role of one's perception of agency. In contrast, research by Reznik et al. (2014) identified an enhanced response in the auditory cortex when participants played simple melodies on a piano keyboard, compared with when these were passively observed. Because participants were required to generate these sounds according to set temporal sequences, it may be argued that enhancement in this context was more likely to have resulted from volitional motor activity than from a higher-level perception of agency. A possible extension to the present study that might help to delineate the effects of ideomotor control from those involving higher-level appraisal of agency may entail providing participants with temporal control over stimuli that are externally triggered. For example, this may be done by allowing participants to select a rhythm or temporal sequence for subsequent passive observation. Further research into the relationship between ideomotor control and perceptions of agency may also have relevance to understanding the pathological substrates of schizophrenia, which is characterised by distortions with respect to both agency (Frith et al., 2000) and sensory attenuation (e.g., Pinheiro et al., 2013; Whitford, 2019).

5. Conclusions

Overall, our findings suggest that the phenomenon of sensory attenuation prevails when controlling for both temporal predictability and temporal control. At the same time, we demonstrate that these factors have differential effects on auditory-evoked activity. Increasing temporal predictability was found to reduce the auditory N1. This is consistent with past research, and suggests that the temporal predictability of self-generated stimuli may in many instances account for an observed reduction in neurophysiological response compared to externally-generated stimuli. Conversely, we observed and describe for the first time (to our knowledge) an apparent amplification of sensory response to stimuli

that are under one’s temporal control. That is, the auditory N1 to stimuli elicited according to participants own timing was found to be larger than when they were required to generate these in response to visual cues. When compared with sensations that are generated by the actions of an external agent, self-generated sensations commonly differ with respect to both temporal predictability and temporal control. Results from the present investigation therefore necessitate a re-evaluation of the experimental paradigms used to study the phenomenon of sensory attenuation.

Data and code availability statement

Data from the present investigation is available via the Open Science Framework and can be accessed at <https://osf.io/rf529/>. This includes processed and segmented EEG data for analysis electrodes, disaggregated by participant. Scripts used for processing and analyses can be downloaded at <https://github.com/a-w-harrison/SA-predictability-control>.

Author contributions

A.W.H, D.J.M, B.N.J, O.G, G.H and T.J.W conceived and designed the experiments. A.W.H. collected and analysed data, and wrote the manuscript under the supervision of T.J.W. and with input from each

of the other co-authors. All authors reviewed and approved the final manuscript.

Declaration of Competing Interest

The authors have no competing interests to declare.

Acknowledgments

We would like to thank the two anonymous reviewers for their insightful and constructive comments, which have helped to improve the quality of this manuscript. This research was supported by the Australian Research Council (ARC) in the form of Discovery Projects awarded to T.J.W. (DP170103094, DP200103288), and the National Health and Medical Research Council of Australia (NHMRC) in the form an Ideas Grant awarded to T.J.W. (APP2004067). A.W.H. is supported by an Australian Government Research Training Program (RTP) scholarship.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.neuroimage.2021.118103](https://doi.org/10.1016/j.neuroimage.2021.118103).

Appendix

Table A1

Paired samples Student’s *t*-tests contrasting N1 amplitudes in Experiment 1.

Comparator 1	M_1	SD_1	Comparator 2	M_2	SD_2	<i>df</i>	M_{diff}	<i>t</i>	<i>p</i>	p_{corr}	<i>d</i>	BF_{10}
Uncued listening	-3.96	1.78	Cued listening	-3.07	2.24	41	-0.89	-3.46	.001	.003	-0.441	24.542
			Uncued self-generation	-3.22	1.93	41	-0.74	-2.60	.013	.019	-0.400	3.219
Cued listening	-3.07	2.24	Cued self-generation	-1.91	2.26	41	-2.05	-6.10	<0.001	<0.001	-1.013	50,782.424
			Uncued self-generation	-3.22	1.93	41	0.15	0.47	.639	.639	0.070	0.185
Uncued self-generation	-3.22	1.93	Cued self-generation	-1.91	2.26	41	-1.16	-2.87	.006	.011	-0.514	5.851
			Cued self-generation	-1.91	2.26	41	-1.31	-3.57	<0.001	.003	-0.624	32.193

Table A2

Paired samples Student’s *t*-tests contrasting P2 amplitudes in Experiment 1.

Comparator 1	M_1	SD_1	Comparator 2	M_2	SD_2	<i>df</i>	M_{diff}	<i>t</i>	<i>p</i>	p_{corr}	<i>d</i>	BF_{10}
Uncued listening	2.18	3.10	Cued listening	2.52	2.94	41	-0.34	-0.81	.423	.462	-0.112	0.227
			Uncued self-generation	-0.14	2.88	41	2.32	4.44	<0.001	<0.001	0.778	352.033
			Cued self-generation	1.73	2.96	41	0.45	0.86	.394	.472	0.148	0.236
Cued listening	2.52	2.94	Uncued self-generation	-0.14	2.88	41	2.66	4.34	<0.001	<0.001	0.915	261.992
			Cued self-generation	1.73	2.96	41	0.79	1.57	.124	.165	0.267	0.517
Uncued self-generation	-0.14	2.88	Cued self-generation	1.73	2.96	41	-1.88	-3.46	.001	.003	-0.643	24.313

Table A3Paired samples Student's *t*-tests contrasting N1 amplitudes in Experiment 2.

Comparator 1	M_1	SD_1	Comparator 2	M_2	SD_2	<i>df</i>	M_{diff}	<i>t</i>	<i>p</i>	p_{corr}	<i>d</i>	BF_{10}
Cued listening	-3.18	1.88	Cued listening (count)	-3.12	2.03	37	-0.06	-0.22	.828	.919	-0.030	0.179
			Uncued self-generation	-3.05	1.61	37	-0.13	-0.33	.746	.933	-0.073	0.184
			Cued self-generation	-2.09	1.89	37	-1.09	-3.08	.004	.023	-0.578	9.343
			Cued self-generation (count)	-2.14	1.57	37	-1.04	-3.09	.004	.029	-0.602	9.496
Cued listening (count)	-3.12	2.03	Uncued self-generation	-3.05	1.61	37	-0.07	-0.16	.873	.903	-0.037	0.177
			Cued self-generation	-2.09	1.89	37	-1.03	-2.85	.007	.031	-0.525	5.555
			Cued self-generation (count)	-2.14	1.57	37	-0.98	-2.47	.018	.055	-0.543	2.499
			Uncued self-generation	-3.05	1.61	Cued self-generation	-2.09	1.89	37	-0.96	-2.70	.010
Uncued self-generation	-3.05	1.61	Cued self-generation (count)	-2.14	1.57	37	-0.91	-3.00	.005	.024	-0.573	7.871
			Cued self-generation (count)	-2.14	1.57	37	0.05	0.18	.856	.917	0.030	0.177

Table A4Paired samples Student's *t*-tests contrasting P2 amplitudes in Experiment 2.

Comparator 1	M_1	SD_1	Comparator 2	M_2	SD_2	<i>df</i>	M_{diff}	<i>t</i>	<i>p</i>	p_{corr}	<i>d</i>	BF_{10}
Cued listening	1.89	2.66	Cued listening (count)	1.78	2.53	37	0.11	0.33	.745	.971	0.042	0.184
			Uncued self-generation	-0.24	2.60	37	2.14	4.48	<0.001	.002	0.812	346.924
			Cued self-generation	1.19	1.87	37	0.71	1.84	.074	.130	0.312	0.802
			Cued self-generation (count)	0.68	2.29	37	1.21	2.30	.027	.058	0.490	1.786
Cued listening (count)	1.78	2.53	Uncued self-generation	-0.24	2.60	37	2.03	4.17	<0.001	.003	0.790	147.793
			Cued self-generation	1.19	1.87	37	0.60	1.80	.081	.135	0.271	0.747
			Cued self-generation (count)	0.68	2.29	37	1.10	2.42	.021	.051	0.457	2.258
			Uncued self-generation	-0.24	2.60	Cued self-generation	1.19	1.87	37	-1.43	-3.88	<0.001
Uncued self-generation	-0.24	2.60	Cued self-generation (count)	0.68	2.29	37	-0.93	-1.92	.063	.117	-0.379	0.913
			Cued self-generation (count)	0.68	2.29	37	0.50	1.77	.085	.127	0.243	0.720

Table A5Paired samples Student's *t*-tests contrasting N2b amplitudes in Experiment 2.

Comparator 1	M_1	SD_1	Comparator 2	M_2	SD_2	<i>df</i>	M_{diff}	<i>t</i>	<i>p</i>	p_{corr}	<i>d</i>	BF_{10}
Cued listening	-0.57	2.58	Cued listening (count)	-1.57	2.81	37	0.99	2.41	.021	.049	0.369	2.203
			Uncued self-generation	-0.61	2.48	37	0.03	0.07	.947	.947	0.013	0.175
			Cued self-generation	-0.75	2.70	37	0.17	0.26	.793	.952	0.066	0.180
			Cued self-generation (count)	-2.23	3.32	37	1.66	2.13	.040	.081	0.563	1.299
Cued listening (count)	-1.57	2.81	Uncued self-generation	-0.61	2.48	37	-0.96	-1.79	.082	.130	-0.364	0.737
			Cued self-generation	-0.75	2.70	37	-0.82	-1.20	.238	.341	-0.298	0.339
			Cued self-generation (count)	-2.23	3.32	37	0.67	0.96	.344	.468	0.218	0.268
			Uncued self-generation	-0.61	2.48	Cued self-generation	-0.75	2.70	37	0.14	0.22	.826
Uncued self-generation	-0.61	2.48	Cued self-generation (count)	-2.23	3.32	37	1.63	2.46	.019	.051	0.562	2.445
			Cued self-generation (count)	-2.23	3.32	37	1.49	2.65	.012	.039	0.495	3.617

References

- Aliu, S.O., Houde, J.F., Nagarajan, S.S., 2009. Motor-induced suppression of the auditory cortex. *J. Cogn. Neurosci.* 21 (4), 791–802. doi:10.1162/jocn.2009.21055.
- Bäfs, P., Jacobsen, T., Schröger, E., 2008. Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *Int. J. Psychophysiol.* 70 (2), 137–143. doi:10.1016/j.ijpsycho.2008.06.005.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc.: Ser. B (Methodol.)* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.
- Blakemore, S.-J., Wolpert, D., Frith, C., 1998. Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635. doi:10.1038/2870.
- Bridgeman, B., 1995. A review of the role of efference copy in sensory and oculomotor control systems. *Ann. Biomed. Eng.* 23 (4), 409–422. doi:10.1007/bf02584441.
- Bruin, K.J., Wijers, A.A., 2002. Inhibition, response mode, and stimulus probability: a comparative event-related potential study. *Clin. Neurophysiol.* 113 (7), 1172–1182. doi:10.1016/S1388-2457(02)00141-4.
- Cohen, J., 1992. A power primer. *Psychol. Bull.* 112 (1), 155. doi:10.1037/0033-2909.112.1.155.
- Donders, F.C. (1969). On the speed of mental processes. In W. G. Koster (Ed.), *Attention and Performance II* (Vol. 30, pp. 412–431). North-Holland. (1868)
- Folstein, J.R., Van Petten, C., 2008. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45 (1), 152–170. doi:10.1111/j.1469-8986.2007.00602.x.
- Ford, J.M., Gray, M., Faustman, W.O., Roach, B.J., Mathalon, D.H., 2007. Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology* 44 (4), 522–529. doi:10.1111/j.1469-8986.2007.00533.x.
- Ford, J.M., Mathalon, D.H., 2012. Anticipating the future: automatic prediction failures in schizophrenia. *Int. J. Psychophysiol.* 83 (2), 232–239. doi:10.1016/j.ijpsycho.2011.09.004.
- Friston, K., 2005. A theory of cortical responses. *Philos.Trans.: Biol. Sci.* 360 (1456), 815–836. doi:10.1098/rstb.2005.1622.
- Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11 (2), 127–138. doi:10.1038/nrn2787.
- Frith, C.D., Blakemore, S.-J., Wolpert, D.M., 2000. Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Res. Rev.* 31 (2), 357–363. doi:10.1016/S0165-0173(99)00052-1.
- Geisser, S., Greenhouse, S.W., 1958. An extension of box's results on the use of the F distribution in multivariate analysis. *Ann. Math. Statist.* 29 (3), 885–891. doi:10.1214/aoms/1177706545.
- Gelman, A., Jakulin, A., Pittau, M.G., Su, Y.-S., 2008. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* 2 (4), 1360–1383. doi:10.1214/08-AOAS191.
- Giard, M.H., Perrin, F., Echallier, J.F., Thévenet, M., Froment, J.C., Pernier, J., 1994. Dissociation of temporal and frontal components in the human auditory N1 wave: a

- scalp current density and dipole model analysis. *Electroencephalogr. Clin. Neurophysiol./Evoked Potentials Sect.* 92 (3), 238–252. doi:10.1016/0168-5597(94)90067-1.
- Gratton, G., Coles, M.G.H., Donchin, E., 1983. A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* 55 (4), 468–484. doi:10.1016/0013-4694(83)90135-9.
- Hackley, S.A., Woldorff, M., Hillyard, S.A., 1990. Cross-modal selective attention effects on retinal, myogenic, brainstem, and cerebral evoked potentials. *Psychophysiology* 27 (2), 195–208. doi:10.1111/j.1469-8986.1990.tb00370.x.
- Haggard, P., 2017. Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18 (4), 196.
- Heinks-Maldonado, T.H., Mathalon, D.H., Gray, M., Ford, J.M., 2005. Fine-tuning of auditory cortex during speech production. *Psychophysiology* 42 (2), 180–190. doi:10.1111/j.1469-8986.2005.00272.x.
- Herrmann, B., Henry, M.J., Scharinger, M., Obleser, J., 2014. Supplementary motor area activations predict individual differences in temporal-change sensitivity and its illusory distortions. *Neuroimage* 101, 370–379. doi:10.1016/j.neuroimage.2014.07.026.
- Herwig, A., Prinz, W., Waszak, F., 2007. Two modes of sensorimotor integration in intention-based and stimulus-based actions. *Q. J. Exp. Psychol.* 60 (11), 1540–1554. doi:10.1080/17470210601119134.
- Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W., 1973. Electrical signs of selective attention in the human brain. *Science* 182 (4108), 177. doi:10.1126/science.182.4108.177.
- Horváth, J., 2015. Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research* 1326, 54–65. doi:10.1016/j.brainres.2015.03.038.
- Horváth, J., Maess, B., Baess, P., Tóth, A., 2012. Action–sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *J. Cogn. Neurosci.* 24 (9), 1919–1931. doi:10.1162/jocn_a.00215/M22360594.
- Houde, J.F., Nagarajan, S.S., Sekihara, K., Merzenich, M.M., 2002. Modulation of the auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14 (8), 1125–1138. doi:10.1162/089892902760807140.
- Hughes, G., Desantis, A., Waszak, F., 2013. Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychol. Bull.* 139 (1), 133–151. doi:10.1037/a0028566.
- Jack, B.N., Le Pelley, M.E., Han, N., Harris, A.W.F., Spencer, K.M., Whitford, T.J., 2019. Inner speech is accompanied by a temporally-precise and content-specific corollary discharge. *Neuroimage* 198, 170–180. doi:10.1016/j.neuroimage.2019.04.038.
- Jeffreys, H., 1998. *The Theory of Probability*. OUP, Oxford.
- Jürgens, U., 1984. The efferent and afferent connections of the supplementary motor area. *Brain Res.* 300 (1), 63–81. doi:10.1016/0006-8993(84)91341-6.
- Klaffehn, A.L., Baess, P., Kunde, W., Pfister, R., 2019. Sensory attenuation prevails when controlling for temporal predictability of self- and externally generated tones. *Neuropsychologia* 132, 107145. doi:10.1016/j.neuropsychologia.2019.107145.
- Lange, K., 2009. Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain Cogn.* 69 (1), 127–137. doi:10.1016/j.bandc.2008.06.004.
- Lange, K., 2011. The reduced N1 to self-generated tones: an effect of temporal predictability? *Psychophysiology* 48 (8), 1088–1095. doi:10.1111/j.1469-8986.2010.01174.x.
- Luck, S.J., Gaspelin, N., 2017. How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology* 54 (1), 146–157. doi:10.1111/psyp.12639.
- Luck, S.J., Hillyard, S.A., 1994. Electrophysiological correlates of feature analysis during visual search. *Psychophysiology* 31 (3), 291–308. doi:10.1111/j.1469-8986.1994.tb02218.x.
- Macar, F., Coull, J., Vidal, F., 2006. The supplementary motor area in motor and perceptual time processing: fMRI studies. *Cogn. Process.* 7 (2), 89–94. doi:10.1007/s10339-005-0025-7.
- Martikainen, M.H., Kaneko, K.-i., Hari, R., 2005. Suppressed responses to self-triggered sounds in the human auditory cortex. *Cereb. Cortex* 15 (3), 299–302. doi:10.1093/cercor/bhh131.
- Mauchly, J.W., 1940. Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Stat.* 11 (2), 204–209. doi:10.1214/aoms/1177731915.
- Mcnorgan, C., 2012. A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery. *Front. Hum. Neurosci.* 6 (285). doi:10.3389/fnhum.2012.00285.
- Mertens, R., Polich, J., 1997. P300 from a single-stimulus paradigm: passive versus active tasks and stimulus modality. *Electroencephalogr. Clin. Neurophysiol./Evoked Potentials Sect.* 104 (6), 488–497. doi:10.1016/S0168-5597(97)00041-5.
- Miall, R.C., Wolpert, D.M., 1996. Forward models for physiological motor control. *Neural Netw.* 9 (8), 1265–1279. doi:10.1016/S0893-6080(96)00035-4.
- Miller, G.A., Gratton, G., Yee, C.M., 1988. Generalized implementation of an eye movement correction procedure. *Psychophysiology* 25 (2), 241–243. doi:10.1111/j.1469-8986.1988.tb00999.x.
- Morey, R., & Rouder, J. (2015). BayesFactor 0.9. 12-2. Comprehensive R Archive Network. In.
- Morillon, B., Baillet, S., 2017. Motor origin of temporal predictions in auditory attention. *Proc. Natl. Acad. Sci.* 114 (42), E8913. doi:10.1073/pnas.1705373114.
- Mulert, C., Jäger, L., Propp, S., Karch, S., Störmann, S., Pogarell, O., Möller, H.-J., Juckel, G., Hegerl, U., 2005. Sound level dependence of the primary auditory cortex: simultaneous measurement with 61-channel EEG and fMRI. *Neuroimage* 28 (1), 49–58. doi:10.1016/j.neuroimage.2005.05.041.
- Näätänen, R., Gaillard, A.W.K., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42 (4), 313–329. doi:10.1016/0001-6918(78)90006-9.
- Näätänen, R., & Picton, T. (1986). N2 and automatic versus controlled processes. In M. W. C. Z. R. & D. F. (Eds.), *Cerebral Psychophysiology: Studies in Event-Related Potentials* (Vol. 38, pp. 169–186). Elsevier.
- Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24 (4), 375–425. doi:10.1111/j.1469-8986.1987.tb00311.x.
- Oestreich, L.K.L., Mifsud, N.G., Ford, J.M., Roach, B.J., Mathalon, D.H., Whitford, T.J., 2015. Subnormal sensory attenuation to self-generated speech in schizotypy: electrophysiological evidence for a 'continuum of psychosis'. *Int. J. Psychophysiol.* 97 (2), 131–138. doi:10.1016/j.ijpsycho.2015.05.014.
- Oestreich, L.K.L., Mifsud, N.G., Ford, J.M., Roach, B.J., Mathalon, D.H., Whitford, T.J., 2016. Cortical suppression to delayed self-initiated auditory stimuli in Schizotypy: neurophysiological evidence for a continuum of Psychosis. *Clin. EEG Neurosci.* 47 (1), 3–10. doi:10.1177/1550059415581708.
- Pinheiro, A.P., Del Re, E., Mezin, J., Nestor, P., Rauber, A., McCarley, R., Gonçalves, Ó.F., Niznikiewicz, M., 2013. Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. *Psychol. Med.* 43 (3), 603. doi:10.1017/s003329171200133x.
- Pinheiro, A.P., Schwartz, M., Gutierrez, F., Kotz, S.A., 2019. When temporal prediction errs: ERP responses to delayed action-feedback onset. *Neuropsychologia* 134, 107200. doi:10.1016/j.neuropsychologia.2019.107200.
- Potts, G.F., Dien, J., Hartry-Speiser, A.L., McDougal, L.M., Tucker, D.M., 1998. Dense sensor array topography of the event-related potential to task-relevant auditory stimuli. *Electroencephalogr. Clin. Neurophysiol.* 106 (5), 444–456. doi:10.1016/S0013-4694(97)00160-0.
- Pritchard, W.S., Shappell, S.A., Brandt, M.E., 1991. Psychophysiology of N200/N400: a review and classification scheme. *Advances in Psychophysiology* 4, 43–106. doi:10.1111/psyp.1201x.
- Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12 (6), 718–724. doi:10.1038/nn.2331.
- Reznik, D., Henkin, Y., Schadel, N., Mukamel, R., 2014. Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nat. Commun.* 5 (1), 4059. doi:10.1038/ncomms5059.
- Sato, A., 2008. Action observation modulates auditory perception of the consequence of others' actions. *Conscious. Cogn.* 17 (4), 1219–1227. doi:10.1016/j.concog.2008.01.003.
- Schafer, E.W.P., Marcus, M.M., 1973. Self-stimulation alters human sensory brain responses. *Science* 181 (4095), 175. doi:10.1126/science.181.4095.175.
- Schwartz, M., Rothermich, K., Schmidt-Kassow, M., Kotz, S.A., 2011. Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biol. Psychol.* 87 (1), 146–151. doi:10.1016/j.biopsycho.2011.02.021.
- Shin, Y.K., Proctor, R.W., Capaldi, E.J., 2010. A review of contemporary ideomotor theory. *Psychol. Bull.* 136 (6), 943–974. doi:10.1037/a0020541.
- Sowman, P.F., Kuusik, A., Johnson, B.W., 2012. Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Exp. Brain Res.* 222 (1), 149–157. doi:10.1007/s00221-012-3204-7.
- Thakkar, K.N., Schall, J.D., Heckers, S., Park, S., 2015. Disrupted Saccadic corollary discharge in Schizophrenia. *J. Neurosci.* 35 (27), 9935–9945. doi:10.1523/jneurosci.0473-15.2015.
- Timm, J., SanMiguel, I., Saupé, K., Schröger, E., 2013. The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neurosci.* 14 (1), 2. doi: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. doi:10.1016/j.cortex.2016.03.018.
- Tran, D.M.D., Harris, J.A., Harris, I.M., Livesey, E.J., 2019. Motor memory: revealing conditioned action tendencies using transcranial magnetic stimulation. *J. Cogn. Neurosci.* 31 (9), 1343–1353. doi:10.1162/jocn_a.01413.
- Tran, D.M.D., Harris, J.A., Harris, I.M., Livesey, E.J., 2020. Motor conflict: revealing involuntary conditioned motor preparation using transcranial magnetic stimulation. *Cereb. Cortex* 30 (4), 2478–2488. doi:10.1093/cercor/bhz253.
- Waszak, F., Cardoso-Leite, P., Hughes, G., 2012. Action effect anticipation: neurophysiological basis and functional consequences. *Neurosci. Biobehav. Rev.* 36 (2), 943–959. doi:10.1016/j.neubiorev.2011.11.004.
- Weiskrantz, L., Elliott, J., Darlington, C., 1971. Preliminary observations on tickling oneself. *Nature* 230 (5296), 598–599. doi:10.1038/230598a0.
- Weiss, C., Herwig, A., Schütz-Bosbach, S., 2011. The self in social interactions: sensory attenuation of auditory action effects is stronger in interactions with others. *PLoS One* 6 (7), e22723. doi:10.1371/journal.pone.0022723.
- Whitford, T.J., 2019. Speaking-induced suppression of the auditory cortex in humans and its relevance to Schizophrenia. *Biol. Psychiatry: Cognitive Neurosci. Neuroimaging* 4 (9), 791–804. doi:10.1016/j.bpsc.2019.05.011.
- Whitford, T.J., Jack, B.N., Pearson, D., Griffiths, O., Luque, D., Harris, A.W.F., Spencer, K.M., Le Pelley, M.E., 2017. Neurophysiological evidence of efference copies to inner speech. *Elife* 6, e28197. doi:10.7554/eLife.28197.
- Wiener, M., Turkeltaub, P., Coslett, H.B., 2010. The image of time: a voxel-wise meta-analysis. *Neuroimage* 49 (2), 1728–1740. doi:10.1016/j.neuroimage.2009.09.064.
- Wolpert, D.M., 1997. Computational approaches to motor control. *Trends Cogn. Sci.* 1 (6), 209–216. doi:10.1016/S1364-6613(97)01070-X.
- Woods, D., 1995. The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalogr. Clin. Neurophysiol.* 44, 102–109. https://www.researchgate.net/publication/15550168_The_component_structure_of_the_N1_wave_of_the_human_auditory_evoked_potential.

- Woods, D., Alho, K., Algazi, A., 1992. Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalogr. Clin. Neurophysiol.* 82 (5), 341–355. doi:[10.1016/0013-4694\(92\)90004-2](https://doi.org/10.1016/0013-4694(92)90004-2).
- Zalta, A., Petkoski, S., Morillon, B., 2020. Natural rhythms of periodic temporal attention. *Nat. Commun.* 11 (1), 1051. doi:[10.1038/s41467-020-14888-8](https://doi.org/10.1038/s41467-020-14888-8).
- Zouridakis, G., Simos, P.G., Papanicolaou, A.C., 1998. Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. *Brain Topogr.* 10 (3), 183–189. doi:[10.1023/a:1022246825461](https://doi.org/10.1023/a:1022246825461).