Title:

Action-effect prediction in sensory attenuation and error monitoring: Distinguishing stimulus-driven and volitional movement

Author names and affiliations:

Anthony W. Harrison[1,*] Bruce K. Christensen,[2] and Thomas J. Whitford[1]

[1] *School of Psychology, University of New South Wales, Sydney, Australia*
[2] *Research School of Psychology, The Australia National University, Canberra, Australia*

Date: 22/08/2022

Key words: Sensory Attenuation; Internal Forward Models; Prediction; Error Monitoring; ERP; Auditory Evoked Potential; Attention; N1; P2; N2b.

[*]**Corresponding author:**

Anthony W. Harrison, School of Psychology, University of New South Wales, UNSW Sydney NSW 2052, anthony.harrison@unsw.edu.au.

Abstract

While volitional movement is believed to be guided by action-effect predictions, these are thought to be less influential with respect to stimulus-driven movement. We compared the effects of predictability on auditory evoked responses to tones that were produced according to participants' ($N = 61$) own timing (i.e., volitionally) and in response to simple visual cues (i.e., stimulus-driven). N2b component amplitudes reflected error monitoring for sound resulting from volitional action, which was absent with respect to stimulus-driven movement. We also explored the sensory attenuation phenomenon, comparing evoked responses to stimuli produced by each form of action with those of externally-generated tones. When controlling for temporal predictability, N1 amplitudes elicited by externally-generated stimuli did not differ from those elicited by volitional or stimulus-driven movement. Reduced P2 amplitudes were observed in response to both volitional and stimulus-driven motor activity, supporting the role of attentional influences in the phenomenon of sensory attenuation.

Introduction

The term 'sensory attenuation' has been used to describe an observed reduction in the subjective intensity and neurophysiological response to stimuli when these are self-generated, compared with those produced externally (e.g., Blakemore et al., 1998; Schafer & Marcus, 1973). This phenomenon is thought to support the prioritisation of externally-generated sensations by suppressing the tides of sensory information associated with an organism's own motor activities while it navigates through its surrounds (see Crapse & Sommer, 2008). Sensory attenuation is believed to be adaptive because, although representation of externally-generated events relies on information arriving through sensory channels, self-generated stimuli may be predicted based on the motor-activities through which they are produced. This notion has contributed to theoretical frameworks suggesting that self-generated sensations are suppressed *based on* predictions associated with eliciting motor commands (see Wolpert, 1997). Such accounts include the influential 'internal forward model' of motor control (IFM; Miall & Wolpert, 1996; Wolpert et al., 1995). According to the IFM, duplicates of motor commands (i.e. 'efference copies'; von Holst & Mittelstaedt, 1950) are transmitted to sensory cortices and activate representations of the anticipated consequences to movement (i.e., 'corollary discharge'; Sperry, 1950). Information arriving through afferent channels are believed to then be compared with these predictions, facilitating suppression of sensations that are anticipated on the basis of the eliciting motor command (Miall & Wolpert, 1996; Wolpert et al., 1995).

In addition to prioritising sensations that result from externally-generated events, computational mechanisms within the IFM are believed to facilitate the selection and enactment of movement based on its predicted outcomes (Wolpert et al., 1995). According to this account, dynamic control is supported through the detection of discrepancies between one's predicted sensations and those arising in response to movement (Miall & Wolpert,

1996). Such discrepancy is represented within this system by residual sensory afference (i.e.,

the portion of one's sensory experience that remains following removal of predictable self-

generated sensations). The notion that action-effect predictions guide movement has been a

central tenet to several other influential theories of motor control, including classical

ideomotor theory (James, 1890; see review by Shin et al., 2010) and more contemporary

perspectives within the active inference literature (Friston, 2009; Friston et al., 2010).

While action-effect prediction is believed to be central to the initiation and control of

some forms of movement, it is considered to be less influential with regard to others.

Previous research has distinguished between movement that is generated on the basis of

action-effect associations (i.e., selection of a motor action based on its intended effects) from

that which involves stimulus-response associations (i.e., movement in reaction to an external

stimulus; Neumann, 1984). Evidence suggests that motor activity may be initiated by either

form of learned association, depending on whether it is volitional (i.e., internally-cued) or

stimulus-driven (i.e., externally-cued; Herwig et al., 2007). In this manner, volitional

movement is believed to be guided by its anticipated sensory consequences (Herwig et al.,

2007; Pfister et al., 2011) while stimulus-driven movement may operate as a form of prepared

reflex with less regard for its specific effects (see Hommel, 2000). In addition to experimental

findings (Herwig et al., 2007), the distinction between volitional and stimulus-driven

movement has been supported by evidence indicating differing neuroanatomical substrates

(see Fried et al., 2017). If stimulus-driven movement is not initiated on the basis of action-

effect predictions or guided by subsequent comparison with resulting sensations, the IFM

does not appear necessary to its operation. Down-weighting of action-effect prediction in the

generation of such movement may leave other proposed functions of the IFM unfulfilled,

however, including prediction-based suppression of self-generated sensations. Suppression of

sensations resulting from stimulus-driven action may therefore involve alternate mechanisms to those supported by action-effect prediction within the IFM.

A distinction between volitional and stimulus-driven action has particular relevance in the context of research that has investigated sensory attenuation through examination of the auditory N1 component (e.g., Han et al., 2021; Han et al., 2022; Schafer & Marcus, 1973; Schröger et al., 2015). The N1 is a large negative component of the auditory evoked potential that occurs between 75 ms and 125 ms after stimulus onset, with subcomponents originating in the supratemporal plane, superior temporal gyrus, and regions within the motor cortex and/or cingulate gyrus (see Giard et al., 1994; Näätänen & Picton, 1987). It is believed to provide a reliable indicator of neurophysiological responses within the primary auditory cortex (see Zouridakis et al., 1998) and has been used extensively in research exploring sensory attenuation (e.g., Schafer & Marcus, 1973; Schröger et al., 2015). In a series of experiments, Harrison et al. (2021) compared the amplitude of N1 components elicited by stimuli that were generated according to participants' own timing (i.e., volitional) and those produced in response to simple visual cues (i.e., stimulus-driven). Results supported an action cueing effect, whereby reduced N1 amplitudes were observed for sounds that were elicited by stimulus-driven action compared with those resulting from volitional movement (Harrison et al., 2021).

The effect of action cueing is particularly important when taken in conjunction with that of temporal predictability, which has been shown separately to influence N1 amplitudes (e.g., Lange, 2009; Weiskrantz et al., 1971) and represents a potential confound in much of the sensory attenuation literature (see Hughes et al., 2013). Significantly, the amplitude of N1 components elicited by sound resulting from volitional action has been found not to differ from that of externally-generated stimuli when these are made predictable in time (Harrison et al., 2021). In contrast, the amplitude of N1 components resulting from sounds produced

through stimulus-driven action have been found to be reduced relative to temporally-predictable externally-generated sounds (Harrison et al., 2021). These findings highlight the importance of controlling for effects relating to temporal predictability in research examining sensory attenuation, and differentiating sensory effects associated with volitional and stimulus-driven action.

Evidence that the auditory N1 is not sensitive to the likelihood of action eliciting sound stimuli (i.e., action-effect contingency) presents another significant challenge to research that has explored IFM-based suppression through examination of this component. Findings from another recent investigation suggest that the amplitude of N1 components elicited by stimuli resulting from volitional and stimulus-driven action are not influenced by action-effect contingency (Harrison et al., 2022). This suggests that suppression of the N1 response to sounds resulting from stimulus-driven action, compared with those produced by either volitional action or externally-generated events, may not be attributable to effects associated with IFM operation. Instead, Harrison et al. (2022) provide tentative evidence that this phenomenon may reflect the outcome of selective attentional mechanisms. Overall, these findings indicate that N1 amplitude is sensitive to the form of action that has elicited stimuli (i.e., whether volitional or stimulus-driven) and that this effect may contribute to differences between self- and externally generated sensations in a manner not associated with motor-based predictions.

While most previous studies of sensory attenuation and other IFM mechanisms have focused on the N1 component, several additional components of the auditory evoked response may provide useful information to understanding these sensorimotor phenomena. This includes the P2 and N2b components in particular. The P2 manifests as a centralised positive component that occurs between 150 ms and 250 ms after stimulus onset (see Crowley & Colrain, 2004). While the functional significance of the P2 remains relatively

poorly understood, several lines of evidence have identified effects involving attentional control. This has included the 'processing negativity', whereby heighted attention to a stimulus contributes to a simultaneous enhancement of its resulting N1 amplitude and a reduction in P2 amplitude (Hillyard et al., 1973; Näätänen et al., 1978). Significantly, an effect consistent with processing negativity has been observed for stimuli produced by volitional action compared with those produced by stimulus-driven action (Harrison et al., 2022).

Attention-related suppression of the P2 has also been observed independently of N1 effects, including with respect to non-target stimuli in auditory oddball tasks when compared with those elicited by identical stimuli when these are not task-relevant (García-Larrea et al., 1992; Novak et al., 1992). These findings have led to the suggestion that reduced P2 amplitude reflects enhanced attentional control during auditory discrimination tasks (Crowley & Colrain, 2004), including processes involved in its termination (García-Larrea et al., 1992). Involvement of the P2 in attentional control is also consistent with evidence demonstrating a linear increase in its amplitude with age, corresponding to an age-related decline in capacity to resist attentional capture (García-Larrea et al., 1992). It has also been noted (see Crowley & Colrain, 2004) that the frontal sites at which these P2 effects are most prominent overlay brain regions involved in protecting against interference by irrelevant stimuli (Dempster, 1992) and those most affected by aging (Pfefferbaum et al., 1998). Examination of P2 components elicited by self- and externally-generated stimuli may therefore serve to further evaluate recent indications regarding involvement of attentional processes in the sensory attenuation phenomenon.

The N2b is a negative frontocentral component that occurs between 200 ms and 350 ms following stimulus onset (see Folstein & Van Petten, 2008). Larger N2b amplitude has been observed in response to stimuli presented during activities with increased cognitive

load, including response inhibition (Bruin & Wijers, 2002) and covert counting tasks

(Harrison et al., 2021; Mertens & Polich, 1997). Significantly, the N2b has also been

implicated in error monitoring processes. In an experiment by Ferdinand et al. (2008),

participants were required to respond to each letter in a sequence with button presses

involving different fingers. A repeating sequence was occasionally interrupted by deviant

trials, in which the presented letter required a response with an alternate finger. Ferdinand et

al. (2008) observed increased N2b amplitudes to self-generated stimuli that deviated from the

sequence, with this effect increasing over the course of the experiment. On this basis, it was

proposed that deviant events were perceived as errors committed by the participant and

reflected in the N2b response. As the IFM framework and other theories of motor control

suggest that predictions regarding the outcome of movement are inherent to its production,

examination of N2b effects is also warranted.

The present study therefore aimed to compare the influence of stimulus predictability

on neurophysiological responses to self- and externally-generated stimuli. We sought to

differentiate sensorimotor processes in stimulus-driven and volitional action, and to assess

and control for the effect of temporal predictability. These aims were achieved through

comparison of the event-related potentials (ERPs) elicited by self- and externally-generated

sound that differed with respect to these factors. Self-generated tones were produced

according to participants' own timing (i.e., volitionally) or in response to simple visual cues

(i.e., stimulus-driven). In the case of externally-generated tones, the absence or presence of

these visual cues supported an assessment of the effects of temporal predictability and its

control in the comparison between self- and externally-generated stimuli (see Fig. 1). To

facilitate comparison of effects relating to action-effect prediction, two variants of each

condition presented sequences of either a single tone type or two different frequencies.

Method

Participants

Sixty-one healthy participants (42 females), aged between 18 and 50 years ($M =$ 22.07, $Mdn = 20.27$, $SD = 5.81$), took part in exchange for credit towards an undergraduate psychology unit at the University of New South Wales (UNSW). The study was approved by the UNSW Human Research Ethic Advisory Panel and recruitment was undertaken through the university's electronic recruitment system (SONA).

Material 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 an EEG recording cap, which was equipped with a BioSemi ActiveTwo system containing 64 Ag/Cl active electrodes. These were positioned according to the extended 10:20 layout. Each participant was also fitted with electrodes below the left eye, on their nose, adjacent to the outer canthus of each eye, and on their left and right mastoids. Participants wore Sennheiser HD201 headphones, placed over the recording cap and electrodes. Grounding was provided by the CMS and DRL cap electrodes while data acquisition was undertaken with a sampling rate of 2048 Hz. In each condition, participants were asked to fix their gaze on a vertical red (fixation) line at the centre of the screen. This line was one pixel wide and approximately six centimetres long (i.e., corresponding to a visual angle of approximately 5.4°).

In *uncued* conditions, a sequence of randomly-spaced white line fragments were presented across five adjacent rows that collectively spanned the height of the fixation line (see Fig. 1B). The line fragments appeared on the right of the screen and moved leftward at a pace of approximately 3°/s. In the *motor-auditory* and *motor* variants of the *uncued* condition, participants were instructed to press a keyboard button repeatedly. They were
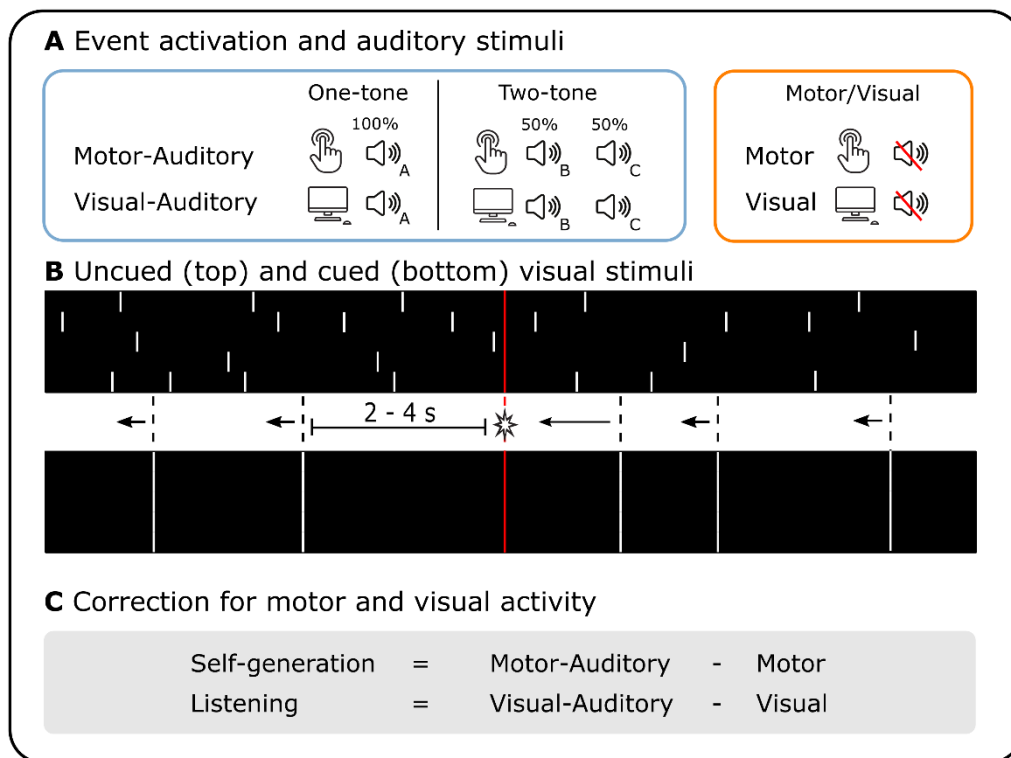
asked to make the timing of their button presses unpredictable to an observer, and to aim for approximately two to four seconds between each press. Brief reminders were displayed above the fixation line from 400 ms to 800 ms post-stimulus following trials in which participants had pressed with less than two seconds between their button presses. The reminder, which asked participants to "slow down", was infrequently required.

In the *uncued motor-auditory* condition, each button press resulted in delivery of a pure tone (85 dB, SPL/A-weighted, 10 ms ramp, 100 ms duration) via participant headphones. In the *one-tone* variant of this condition, all tones were of the same frequency. These could be either *low*, *medium* or *high* pitch tones, corresponding to 500 Hz, 1000 Hz and 1500 Hz respectively. In a *two-tone* variant, each button press elicited one of two different tones. The total number of each tone type was balanced within blocks and presented in randomised order. Allocation of the three tone frequencies to the *one-tone* and *two-tone* variants was counterbalanced across participants. In the *motor* variant of the *uncued* condition, silent audio tracks were triggered to mark each button press. Latency of audio track delivery (i.e., both tones and silent tracks) was reduced with the assistance of an AudioFile Stimulus Processor (Cambridge Research Systems).

Participants observed passively in the *visual-auditory* and *visual* variants of the *uncued* condition. In the *one-tone* variant of the *visual-auditory* condition, a sequence of tones with a single frequency was presented to participants. In the *two-tone* variant, a sequence of two different tone frequencies were again presented. The allocation of tone frequencies to each condition, and their order within *two-tone* sequences, was the same as for the *uncued motor-auditory* conditions. The timing of their presentation was also matched to the timing of participants' button presses in the preceding *uncued motor-auditory* condition. In *visual* conditions, the timing of silent audio tracks was determined based on this approach.

In *cued* conditions, vertical white (stimulus) lines appeared on the right of the screen and moved leftward at a constant rate of approximately 3°/s. These lines were one pixel wide and approximately six centimetres long (i.e., equal dimensions to the fixation line; see Fig. 1B). In the *motor-auditory* and *motor* variants of the *cued* condition, participants were instructed to press the keyboard button at the precise moment that each stimulus line intersected with the fixation line. The spacing of the lines, and therefore the timing of events, was based on the spacing of participants' button presses in the preceding *uncued motor-auditory* or *uncued motor* block (i.e., the most recently presented). Button presses in the *cued motor-auditory* condition elicited the same auditory stimuli as described for the *uncued motor-auditory* and *uncued visual-auditory* conditions, with different blocks presenting *one-tone* and *two-tone* sequences. Button presses again triggered delivery of a silent audio track to mark events in the *cued motor* condition. Participants were instructed to passively observe in the *cued visual-auditory* and *cued visual* conditions. In the *cued visual-auditory* condition, tones were presented at the precise moment that each stimulus line intersected with the fixation line. Auditory stimuli were the same as described for other conditions, with *one-tone* and *two-tone* block varieties. Silent audio tracks were used to mark these events in the *cued visual* condition.

Participants completed a single block of each of the twelve experimental conditions. Each block involved a total of 70 trials. The order of blocks was pseudorandomised, such that the *uncued motor-auditory* and *uncued motor* conditions were presented in the first or seventh block. The allocation of these conditions to each block was counterbalanced across participants, while the order of the 10 remaining conditions was randomised for each participant.

Fig. 1: Protocol schematic and visual stimuli

A Event activation and auditory stimuli

B Uncued (top) and cued (bottom) visual stimuli

2 - 4 s

C Correction for motor and visual activity

Self-generation	=	Motor-Auditory	-	Motor
Listening	=	Visual-Auditory	-	Visual

A. In *motor-auditory* and *motor* conditions, participants pressed a keyboard button to initiate events. These were initiated externally, by the computer, in the *visual-auditory* and *visual* conditions. Events involved presentation of a tone via participant headphones in the *motor-auditory* and *visual-auditory* conditions, while silent audio tracks were used to mark events in the *motor* and *visual* conditions. In *one-tone* variants of the *motor-auditory* and *visual-auditory* conditions, all tones were of a single frequency. A balanced and randomised sequence two different tone frequencies were presented in the *two-tone* variants. **B.** In *uncued* conditions, vertical line fragments were randomly distributed across five adjacent rows and moved leftwards across the screen. In the *motor-auditory* and *motor* variant, participants were instructed to ignore these lines, keep their eyes fixed on a red fixation line and press a keyboard button once every two to four seconds approximately. The timing of events (i.e., inter-trial intervals) in the preceding *motor-auditory* condition provided the bases for timing in subsequent *visual-auditory* and *visual* conditions. In *cued*

conditions, this timing was represented by longer vertical lines that moved from right to left across the screen. In *motor-auditory* and *motor* variants of the *cued* condition, participants were instructed to press the keyboard button to initiate events at the precise moment that each line intersected with the fixation line. Events were externally triggered at this precise in the *visual-auditory* and *visual* variants. **C.** *Motor-auditory* conditions were corrected for motor (and visual) activity by subtracting the equivalent *motor* condition (i.e., *uncued* and *cued*). Similarly, visual activity was removed from the *visual-auditory* conditions by subtracting the equivalent *visual* condition. Motor-corrected motor-auditory conditions are described as *self-generation,* while visual-corrected visual-auditory conditions are described as *listening.* Twelve experimental conditions were thereby used to develop eight analysis conditions, including *one-tone* and *two-tone* variants of the *uncued self-generation, cued self-generation, uncued listening* and *cued listening* conditions.

EEG Processing

EEG data were referenced to the nose electrode and processed offline in BrainVision Analyser (Version 2.2.0; Brain Products GmbH, Gilching, Germany). A 50-Hz notch filter and 0.1 Hz to 30 Hz zero phase shift Butterworth band-pass filter (half-amplitude, high-pass 12 dB/Oct and low-pass 48 dB/Oct slopes) were applied. Trials were segmented into 600 ms epochs, commencing 200 ms prior to each event. Recordings from the electrodes below participants' left eyes were used in conjunction with those at Fp1 to produce vertical electro-oculograms (EOGs). Horizontal EOGs were similarly constructed based on recordings at the electrodes that were placed adjacent to the outer canthus of each eye. EOGs were then used to correct eye movement artefacts using the method described by Miller et al. (1988), which is an extension to the approach by Gratton et al. (1983). Baseline correction was applied using average recordings in the 200 ms prior to each event. Epochs that were found to contain

peak-to-peak amplitudes in excess of 200 µV were then excluded. Averaged waveforms were developed by condition and participant, based on the remaining epochs.

Activity associated with the auditory stimuli in *motor-auditory* and *visual-auditory* conditions was isolated by subtracting equivalent *motor* and *visual* waveforms. Twelve experimental conditions thereby supported calculation of eight analysis conditions, including *one-tone* and *two-tone* varieties for each task conditions, including *uncued self-generation, cued self-generation, uncued listening* and *cued listening.*

Consistent with previous analyses (e.g., Harrison et al., 2021), the N1 and N2b were analysed based on pooled recordings at Fz, FCz and Cz, while the P2 was analysed based on pooled recordings at FCz, Cz and CPz. Peak latencies of the N1 and P2 components were identified using the collapsed localizer method (see Luck & Gaspelin, 2017). This involved identifying component latency within a collapsed waveform that represented the average voltages across all participants and conditions. When defined as the most negative local minimum between 50 ms and 150 ms, the N1 was found to have occurred 90.3 ms post-auditory stimulus. The P2 was found to have occurred 161.1 ms after stimulus onset, when defined as the most positive local maximum between 40 ms and 170 ms post-auditory. Because a clear N2b component was only evident in the *two-tone uncued self-generation* condition (see Fig. 2), the grand-averaged waveform of this condition was used as the basis for determining its latency. When taken as the most negative local minimum between 150 ms and 300 ms post-auditory, the N2b peak was found to have occurred 200.7 ms post-auditory stimulus. An apparent P3 component was also observed in the *two-tone uncued self-generation* condition, with a latency of 279.3 ms when taken as the most positive local maximum between 250 ms and 350 ms.

Statistical Analyses

Statistical analyses were based on average recordings, by condition and participant, within 20 ms windows that were centred on the peak latencies identified above. Analyses of the N1 involved average recordings between 81 ms and 101 ms post-auditory stimulus, while the N2b was analysed based on those between 179 ms and 199 ms post-auditory stimulus.

Three 2 x 4 repeated measures ANOVAs were used to investigate the main effects of *sequence* (i.e., *one-tone* versus *two-tone*) and *task*, as well as their interaction, on N1, P2 and N2b amplitudes respectively. Planned comparisons involved three orthogonal contrasts of condition tasks (i.e., *uncued self-generation, cued self-generation, uncued listening* and *cued listening)*, supporting investigation of the effects of *volition* ([3, -1, -1, -1]), *self-generation* ([0, 2, -1, -1]) and *temporal predictability* ([0, 0, -1, 1]). Bayes factors (BF_{10}) were also produced for the omnibus repeated measures ANOVAs and planned comparisons. These included Cauchy priors with an r-scale of $1/\sqrt{2}$ (see Morey & Rouder, 2018), as recommended across a range of Bayesian statistical procedures including point null hypothesis testing (Jeffreys, 1998) and regression (Gelman et al., 2008).

With regard to behavioural data, consistency in inter-trial interval (ITI) was assessed across experimental conditions based on a 3 x 2 x 2 repeated measures ANOVA. In particular, this investigated differences based on the nature of *auditory stimuli* (i.e., *motor/visual, one-tone* and *two-tone*), *event activation* (i.e., *self-generation* versus *external*) and *visual stimuli* (i.e., *uncued* versus *cued*).

Results

N1

A 2 x 4 repeated measures ANOVA was used to investigate the main effects of *sequence* (i.e., *one-tone* versus *two-tone*) and *task*, as well as their interaction, on N1 component amplitudes (see *Statistical Analyses)*. Planned comparisons involved three orthogonal contrasts of condition *task* (i.e., *uncued self-generation, cued self-generation, uncued listening* and *cued listening)*, supporting investigation of the effects of *volition* ([3, -1, -1, -1]), *self-generation* ([0, 2, -1, -1]) and *temporal predictability* ([0, 0, -1, 1]). Non-sphericity was identified by Mauchly's test (Mauchly, 1940) and corrected using the Greenhouse-Geisser method (Geisser & Greenhouse, 1958). The ANOVA revealed statistically significant main effects of both *sequence*, $F(0.86, 51.72) = 11.86$, $p = .001$, $\eta_p^2 = 0.17$, $BF_{10} = 9.18$, and *task*, $F(2.94, 176.56) = 2.91$, $p = .036$, $\eta_p^2 = 0.05$, $BF_{10} = 1.76$ (see Fig. 2). Planned comparisons indicated a significant effect of *temporal predictability* (i.e., contrasting *cued listening* and *uncued listening)*, $t(180) = 2.34$, $p = .020$, $d = 0.30$, $BF_{10} = 33.62$, such that *cued listening* exhibited a smaller (i.e., less negative) N1 amplitude than *uncued listening*. The effects of *volition* and *self-generation* were not found to be significant (see Table A3 for detail). Nor were interactions between the effect of *sequence* and each of the planned comparisons. This included the *temporal predictability* x *sequence* interaction, $t(240) = -0.32$, $p = .751$, $d = -0.04$, $BF_{10} = 0.11$, and the *self-generation* x *sequence* interaction, , $t(240) = -0.13$, $p = .896$, $d = -0.02$, $BF_{10} = 0.17$. In both cases, the associated Bayes factor indicated substantial evidence in favour of the null hypothesis (see Jeffreys, 1998).
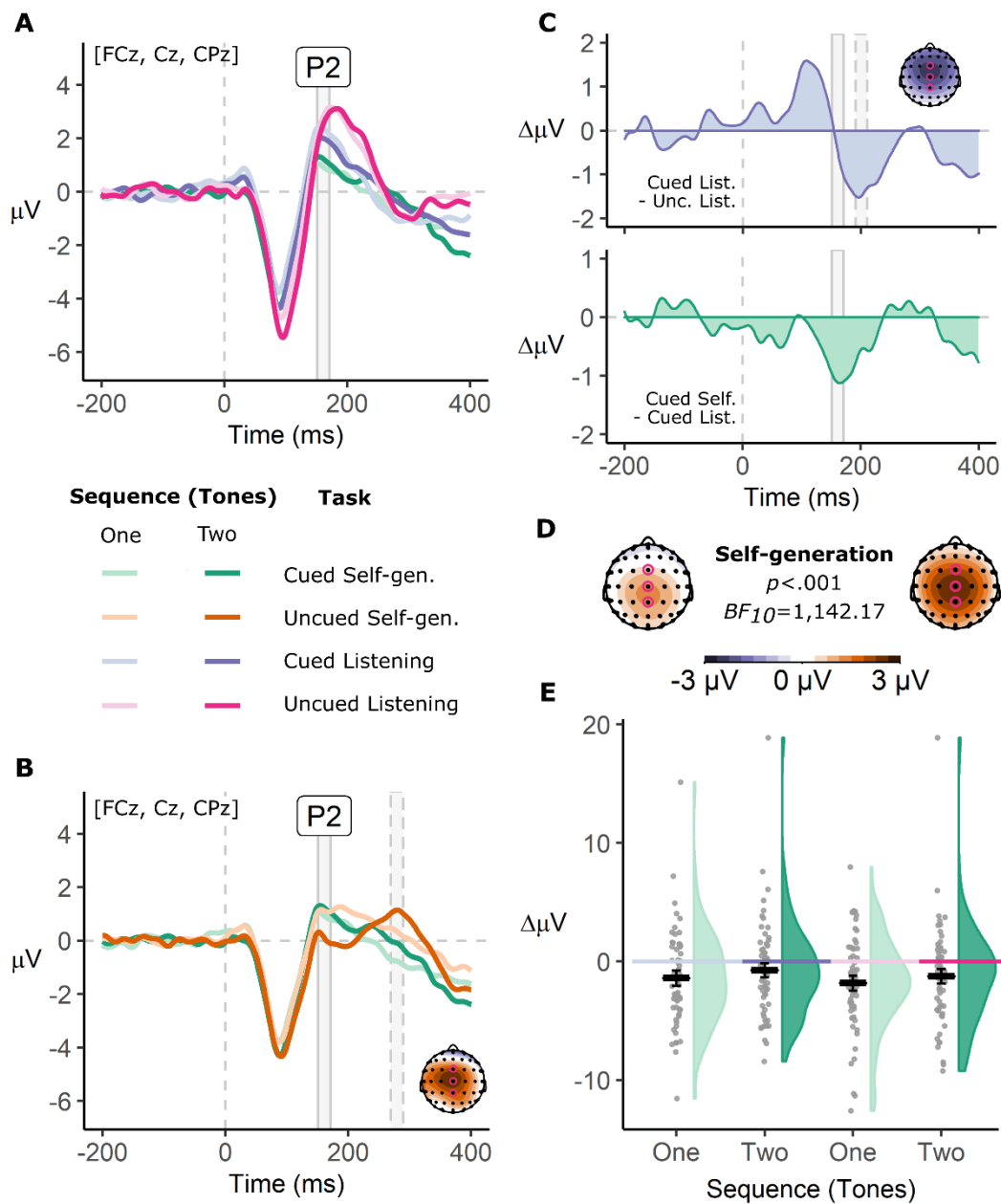
Fig. 2: Analyses of N1 component amplitudes
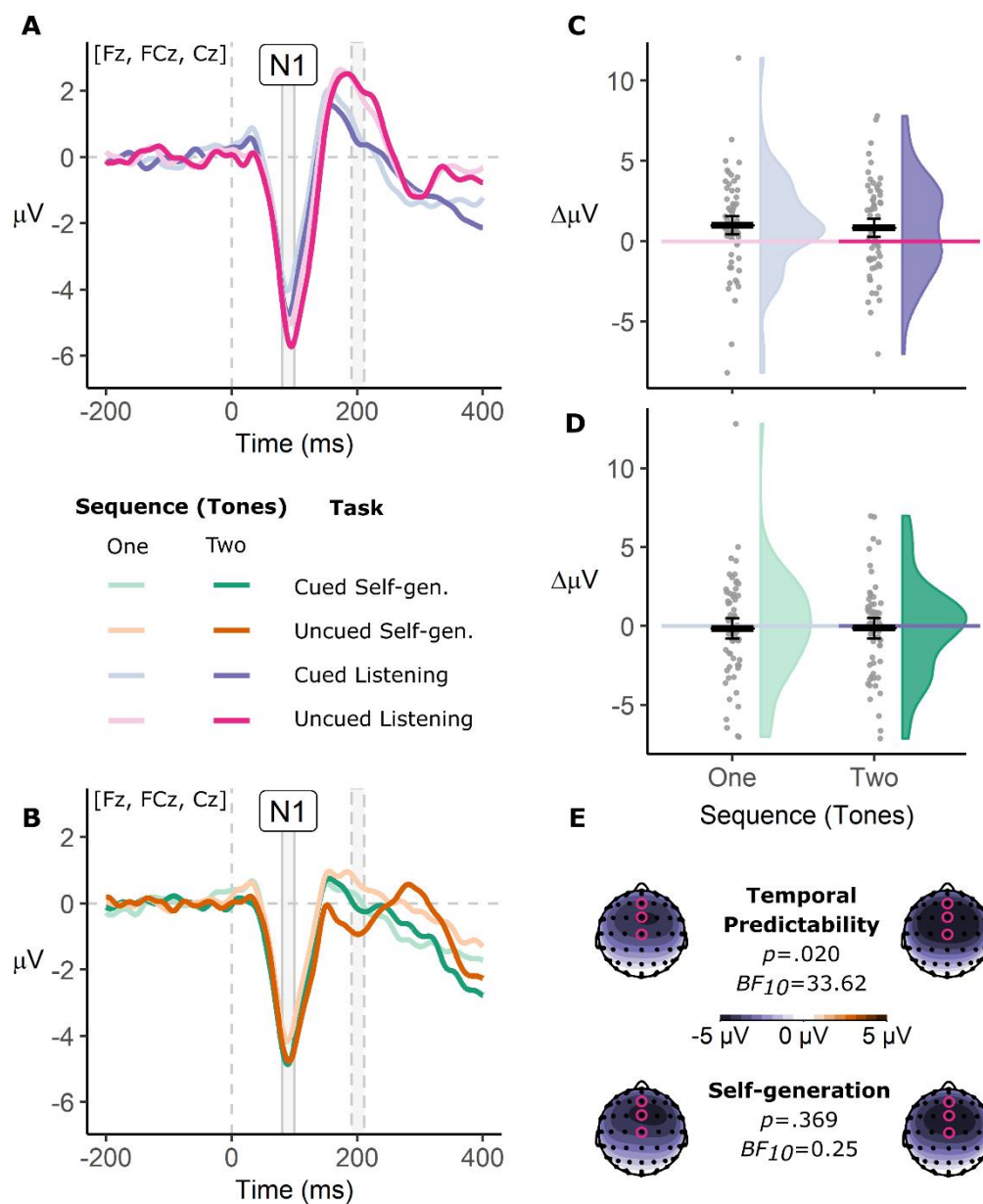
A-B. Pooled grand-averaged recordings at Fz, FCz and Cz, demonstrating mean amplitude

for (A) *listening* conditions and (B) *self-generation* conditions, by *sequence* (i.e., *one-tone*

and *two-tone*). Dotted rectangles mark N2b analysis window. **C-D.** Within-subject

differences in mean N1 amplitude between (C) *cued listening* and *uncued listening* and (D)

cued self-generation and *cued listening,* by *sequence* (i.e., *one-tone* and *two-tone*). **E.**

Statistical test outcomes relating to the effects of temporal predictability and self-

generation. Accompanying topographic maps demonstrate the contrast in N1 distribution

between the (top left) *cued listening* and (top right) *uncued listening* conditions, and

between (bottom left) *cued self-generation* and combined (bottom right) *listening*

conditions. Pink rings indicate analysis electrodes.

P2

Differences in P2 component amplitude were also investigated using a 2 x 4 repeated

measures ANOVA. This was used to examine effects involving experimental *sequence* and

task, as well as their interaction. Non-sphericity was again corrected using the Greenhouse-

Geisser method (Geisser & Greenhouse, 1958). The ANOVA revealed a statistically

significant effect of *task* on P2 amplitude, $F(2.22, 133.15) = 9.25$, $p < .001$, $\eta_p^2 = 0.13$, $BF_{10} =$

334,906.58 (see Fig. 3). In contrast, non-significant effects were observed with respect to

sequence, $F(0.77, 45.95) = 1.67$, $p = .201$, $\eta_p^2 = 0.03$, $BF_{10} = 0.23$, and the interaction

between *task* and *sequence*, $F(2.30, 137.84) = 1.97$, $p = .120$, $\eta_p^2 = 0.03$, $BF_{10} = 0.07$. Planned

comparisons revealed statistically significant effects of both *self-generation*, $t(180) = -3.57$, p

$< .001$, $d = -0.46$, $BF_{10} = 1,142.17$, and *volition*, $t(180) = -2.09$, $p = .038$, $d = -0.31$, $BF_{10} =$

1,223.77, on P2 amplitude. Significance of the *self-generation* effect suggested that the *cued*

self-generation condition had a smaller (i.e., less positive) P2 amplitude than the *listening*

conditions (i.e., *uncued listening* and *cued listening)*. The nature of the *volition* effect was

such that the *uncued self-generation* condition exhibited a smaller P2 amplitude than the *cued*

self-generation condition. Other planned comparisons and interactions with the effects of

sequence, were not found to be statistically significant (see Table A3).

Fig. 3: Analyses of P2 component amplitudes

A-B. Pooled grand-averaged recordings at FCz, Cz and CPz, demonstrating mean

amplitude for (A) *cued self-generation, cued listening* and *uncued listening* conditions and

(B) all *self-generation* conditions, by *sequence* (i.e., *one-tone* and *two-tone*). Dotted

rectangle and accompanying topographic map in (B) demonstrate the P3 component (279

ms post-stimulus) in the *two-tone uncued self-generation* condition. **C.** Difference waves,

representing (top) *cued listening* minus *uncued listening,* and (bottom) *cued self-generation*

minus *cued listening* (i.e., combined one-tone and two-tone variants). Difference waves are collapsed across sequence (i.e., combining one- and two-tone variants). Dotted rectangle and accompanying topographic voltage map demonstrate the distribution of difference waves at the N2b analysis window, appearing to reflect reduced latency and amplitude of the P2 in *cued listening* compared with *uncued listening*. **D.** Statistical test outcomes relating to the effect of self-generation and accompanying topographic maps, including (left) *cued self-generation* and (right) combined *listening* conditions. Pink rings indicate analysis electrodes. **E.** Within-subject differences in mean P2 amplitude, including *cued self-generation* minus *cued listening* and *uncued listening* conditions by *sequence* (i.e., *one-tone* and *two-tone*).

N2b

To investigate effects involving the N2b component, another 2 x 4 repeated measures ANOVA was conducted. This again supported examination of the main effects of *sequence* and *task*, as well as their interaction. Mauchly's test identified non-sphericity (Mauchly, 1940), which was again corrected using the Greenhouse-Geisser method (Geisser & Greenhouse, 1958). Results indicated a statistically significant effect of *task, F*(2.46, 147.33) = 10.68, $p < .001$, $\eta_p^2 = 0.15$, $BF_{10} = 4,336,194.87$. While the main effect of sequence was not significant, F(0.80, 47.73) = 2.14, $p = .149$, $\eta_p^2 = 0.03$, $BF_{10} = 0.26$, a statistically significant interaction was observed between *task* and *sequence*, F(2.39, 143.18) = 4.71, $p = .003$, $\eta_p^2 = 0.07$, $BF_{10} = 0.42$ (see Fig. 4). Planned comparisons revealed significant effects associated with *self-generation, t*(180) = -2.41, $p = .017$, $d = -0.31$, $BF_{10} = 1,254.35,$ and *temporal predictability*, t(180) = -2.23, $p = .027$, $d = -0.29$, $BF_{10} = 6,692.37$. The effect of *self-generation* was such that the *cued self-generation* condition had larger (i.e., more negative) N2b components compared with those of the *listening* conditions (i.e., *uncued listening* and

cued listening). The temporal predictably effect was such that larger N2b components were observed in the *cued self-generation* condition compared with the *uncued self-generation* condition. Interestingly, a significant interaction was also observed between *volition* and *sequence, t*(240) = -3.28, *p* = 0.001, *d* = -0.49, *BF_{10}* = 3.57. This interaction was such that the *two-tone* sequence resulted in larger N2b amplitudes than the *one-tone* sequence in the *uncued self-generation* condition, while this effect was not apparent with respect to the *cued self-generation* condition.

Fig. 4: Analyses of N2b component amplitudes
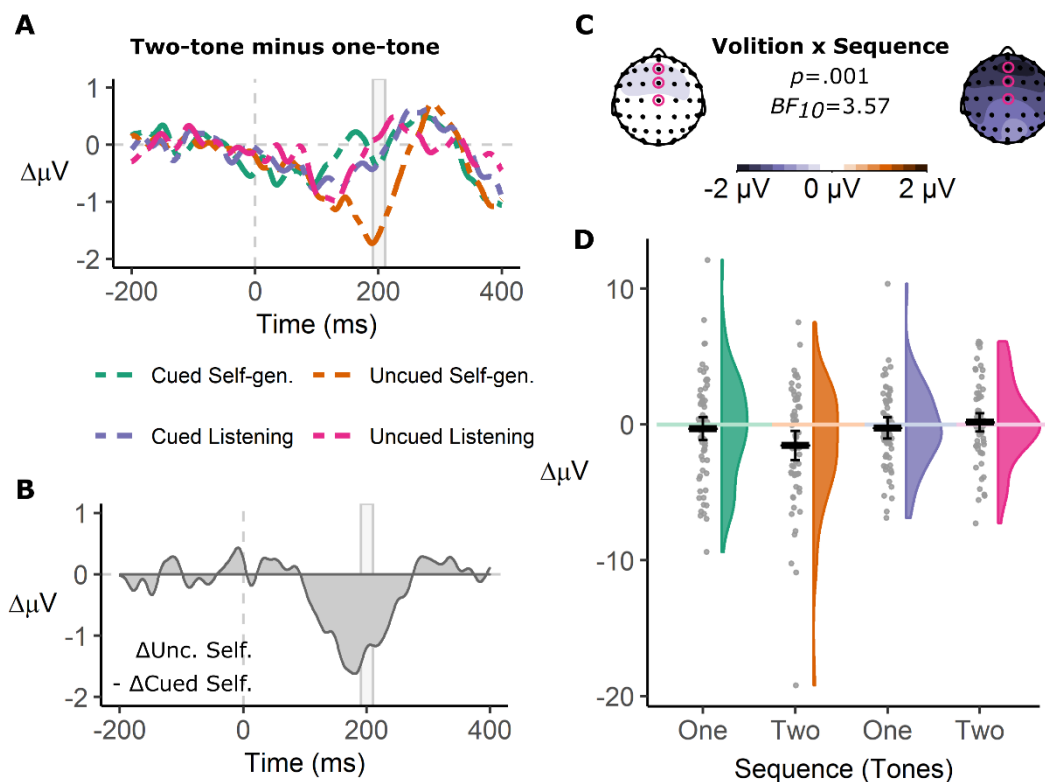
A. Difference waves representing *two-tone* minus *one-tone* variants by task condition, including *cued self-generation, uncued self-generation, cued listening* and *uncued listening* condition. **B.** Difference wave, representing differences in the effect of sequence (i.e., *two-tone* minus *one-tone*) between *cued self-generation* and *uncued self-generation* conditions. **C.** Statistical test outcomes relating to the interaction between volition and sequence.

Accompanying topographic maps represent the effect of sequence (i.e., *two-tone* minus *one-tone*) for (left) combined *cued self-generation, cued listening* and *uncued* listening, and (right) *uncued self-generation.* Pink rings indicate analysis electrodes. **D.** Within-subject differences in mean N2b amplitude by sequence (i.e., *two-tone* minus *one-tone*) for each task condition, including *cued self-generation, uncued self-generation, cued listening* and *uncued listening.*

Behavioral

Descriptive statistics relating to ITI are provided by condition in Table A4. A 3 x 4 repeated measures ANOVA was used to assess consistency in ITI across experimental conditions. Specifically, this examined differences based on the nature of *auditory stimuli* (i.e., motor/visual, one-tone and two-tone), *event activation* (i.e., self-generated versus external) and *visual stimuli* (i.e., uncued versus cued). Following correction for non-sphericity using the Greenhouse-Geisser method (Geisser & Greenhouse, 1958), results from the ANOVA identified non-significant effects with respect to each factor, including *auditory stimuli, F*(1.96, 117.60) = 0.07, *p* = .930, η_p^2 = 0.00, *BF$_{10}$* = 0.02, *event activation, F*(0.84, 50.70) = 0.35, *p* = .554, η_p^2 = 0.01, *BF$_{10}$* = 0.10, and *visual stimuli, F*(0.94, 56.39) = 1.81, *p* = .184, η_p^2 = 0.03, *BF$_{10}$* = 0.32. Interactions were also found to be non-significant (see Table A5 for detail). Overall, these results reflect relative consistency in ITI across experimental conditions.

Power Analyses

The power (1 - β) of the sample (*N* = 61) to assess small, medium, and large effect sizes, was explored through post hoc analyses. Based on standardised reporting conventions (Cohen, 1988), the sample was sufficient to detect small (η_p^2 = .01), medium (η_p^2 = .06) and

large (η_p^2 = .14) two-level repeated measures main effects with powers of .339, .973, and <.999, respectively. With regard to three-level main effects, the sample was found to achieve powers of .379, .993, and <.999, respectively.

Discussion

In the present investigation, we compared the effects of predictability on evoked responses to self- and externally-generated stimuli. The study was designed to evaluate and control for effects relating to differences in temporal predictability, serving to address prior methodological limitations in this regard. We also differentiated between stimulus-driven and volitional movement, given evidence demonstrating distinct neuroanatomical substrates and sensorimotor processing. Importantly, the study sought to provide an exploratory analysis of relevant components across the evoked response. In addition to the auditory N1, this included analyses of P2 and N2b component amplitudes.

Findings replicated previous research (e.g., Lange, 2009; Weiskrantz et al., 1971) in demonstrating reduced N1 amplitudes (i.e., relative positivity) to stimuli when these were made predictable in time (i.e., *cued listening* versus *uncued listening)*. The effect of *self-generation* was not found to be significant when controlling for *temporal predictability*, with N1 amplitudes in the *cued listening* condition not found to differ from those relating to self-generated stimuli in the *uncued self-generation* and *cued self-generation* conditions. The corresponding Bayes Factor suggests substantial evidence in favour of mean amplitudes not differing in this regard. This finding is in contrast with previous research that has identified differences between the amplitude of N1 components elicited by self- and externally-generated stimuli. In this manner, they provide further indication that such differences may in many instances be attributed to the inherent temporal predictability of self-generated sensations.

In contrast with recent findings (Harrison et al., 2022; Harrison et al., 2021), the effect of *volition* (i.e., contrasting *cued self-generation* with *uncued self-generation* and both *listening* conditions) was not found to be significant either. Although the reasons for this difference are not immediately apparent, it is noteworthy that the current study involved longer ITIs than in previous investigations. This was due to the inclusion of prompts, which reminded participants to slow down following trials with an ITI less than 2000 ms. Future research may therefore examine the potential that the effect of action cueing on N1 is moderated the interval between self-generated stimuli.

The effect of *sequence* (i.e., contrasting *one-tone* and *two-tone* variants of each condition) on N1 amplitude was significant, reflecting a reduced primary cortical response in series of trials involving only one tone frequency. However, the non-significant *sequence* x *self-generation* interaction suggested that N1 amplitudes were not differentially affected by tone frequency prediction for self- and externally-generated stimuli. This finding is consistent with evidence that N1 amplitude is not influenced by the likelihood of action eliciting a specific stimulus (Harrison et al., 2022), and is contrary to the attribution of reduced N1 amplitudes for self-generated stimuli to the effects of motor-based prediction. The *sequence* x *temporal predictability* interaction was also not found to be significant, suggesting that tone frequency prediction did not confer additional suppression to passively observed stimuli when these were made predictable in time. A significant main effect of *sequence* may therefore be consistent with the effects of repetition suppression, whereby a reduction in neural response is observed upon repeated presentation of a stimulus (Desimone, 1996). This effect is believed to reflect the refined engagement of neural populations (Wiggs & Martin, 1998) and desensitisation through neuronal fatigue (Grill-Spector et al., 2006).

Self-generated stimuli were found to produce significantly smaller P2 amplitudes compared to those resulting from externally-generated events. Although the functional

significance of the P2 remains relatively poorly understood, several lines of evidence have

implicated the component in attentional processes. This includes the well-established

'processing negativity' effect (Hillyard et al., 1973), in which heightened attention to stimuli

results in a negativity that spans the N1 and P2 range. Attention-related suppression of the P2

has also been observed independently of N1 effects, as an apparent result of task relevance

(García-Larrea et al., 1992; Novak et al., 1992) and alertness (Colrain et al., 2000). In

particular, it has been proposed that suppression of the P2 may reflect an executive control

process supporting withdrawal of attention from non-relevant stimuli (García-Larrea et al.,

1992). Reduced P2 amplitude for self-generated stimuli may therefore suggest that these were

subject to heightened attentional control compared with levels allocated to stimuli produced

by externally-generated events. This finding supports the notion that sensory attenuation may

be partly attributable to effects involving attention, including its withdrawal from non-

relevant stimuli as indexed by reduced P2 amplitude.

Interestingly, the suppressive effect of temporal predictability on N1 amplitude (i.e.,

cued listening versus *uncued listening)* was followed by relative negativity that was

statistically significant at its peak approximately 200 ms post-stimulus (i.e., the N2b analysis

window). This was not followed by an apparent P3 and was therefore unlikely to reflect

influences associated with the N2b component, as the N2b requires attention to eliciting

stimuli and is always accompanied by a P3 (Folstein & Van Petten, 2008; Pritchard et al.,

1991). Instead, it may also be attributable to attentional modulation of the P2, which has been

found to be larger and occur later in lower states of arousal, such as when transitioning from

wakefulness to sleep (Colrain et al., 2000). Reduced P2 amplitudes in the temporally-

predictable *cued listening* condition may therefore reflect greater preparedness for withdrawal

of attention to cued stimuli. That this suppression occurred most prominently at frontal sites

(see Fig. 3C) is also consistent with the effects of attention that have previously been observed with regard to the P2 component (García-Larrea et al., 1992; Novak et al., 1992).

Findings in relation to the N2b have particular relevance to understanding the role of action-effect prediction in motor control. A significant interaction was observed between *sequence* and *volition* (i.e., *uncued self-generation* in comparison with *cued self-generation* and both *listening* conditions), involving a prominent N2b in the *two-tone* variant of the *uncued self-generation* task that was not apparent in the *one-tone* variety (see Fig. 4). In contrast, N2b components were not evident in either variant of the *cued self-generation, cued listening* or *uncued listening* conditions. Research has implicated the N2b in error monitoring processes relating to self-generated sensations, with evidence to suggest that larger component amplitudes occur in response to unexpected task-relevant stimuli (Ferdinand et al., 2008; Ferdinand et al., 2015). This finding may therefore be consistent with the notion that volitional action involves comparison of the intended and actual outcomes to movement, while stimulus-driven action may be enacted with less regard for its specific sensory effects. In conjunction with recent evidence demonstrating differences in sensorimotor processing, this result highlights the need to distinguish between these forms of movement in research examining neurophysiological responsivity to self-generated stimuli.

Our results may also have relevance to understanding the mechanisms that give rise to one's sense of agency. Because the selection of a motor action is based on an agent's expectations regarding its capacity to achieve intended outcomes, sense of agency is believed to also relate to the alignment of predicted and actual sensory experience (see Haggard, 2017). Evidence for this relationship has been provided by research into the phenomenon of intentional binding, whereby the perceived interval between deliberate action and subsequent stimulus is reduced (Haggard et al., 2002). Moore and Haggard (2008) found that, for an action that does not consistently result in the presentation of a stimulus, temporal binding of

the action (i.e., the degree to which it is experienced as having occurred later in time) is larger

when stimuli are subsequently presented than when they are not. This suggests that

intentional binding may retrospectively affect the perceived timing of action, based on

resulting sensations. Significantly, for action not resulting in the presentation of a stimulus,

Moore and Haggard (2008) found that intentional binding was larger when the prospective

likelihood of action eliciting a stimulus was higher. Intentional binding thereby appears to be

influenced by both prospective and retrospective factors, which contribute to one's sense of

agency through the prediction of action-effects and their observation. Our findings suggest

that, in contrast to volitional action, the processing of sounds resulting from stimulus-driven

action is not influenced by predictions regarding its effects. For this reason, the retrospective

aspect to one's assessment of agency may not be supported for this form of movement.

Future research may seek to explore this possibility.

The study contained several limitations that warrant acknowledgment. Firstly, clear

challenges are present with regard to disentangling effects relating to components with

overlapping latencies. Future research may seek to distinguish effects involving the P2 and

N2b in particular, given the close proximity of these components. Difficulty associated with

the distinction of these components was most evident in the interpretation of effects

associated with *temporal predictability*. Although the *cued listening* conditions were found to

be more negative than the *uncued listening* conditions in the N2b analysis window (see Fig.

3C), this effect was attributed to differences in the P2 as this was supported by a stronger

theoretical basis. This included the likelihood of differences in attentional control, and the

absence of an apparent P3 in the *cued listening* condition. A second limitation relates to the

presentation of trials for each condition type in separate blocks, which meant that overall

levels of arousal may not necessarily be assumed to be equal. Although amplitudes of each

analysed component may have been subject to arousal effects, a lack of consistency in the

observed influence could serve to disconfirm this potential. For example, while the larger P2 component amplitudes for the *uncued listening* condition would indicate a general reduction in arousal, this is not consistent with larger N1 amplitudes observed for the same condition. Nevertheless, future research may seek to investigate the observed effects in a design that better controlled for potential differences in attention and arousal.

Conclusions

Findings from the present investigation highlight differences in the way that motor-based sensory prediction influences the processing of sound produced through volitional and stimulus-driven action. Self-generated tones with unpredictable sound frequencies were found to elicited N2b components when these were produced through volitional action, while self-generated tones with predictable sound frequencies did not. In contrast, N2b components did not appear in response to either predictable or unpredictable tone frequencies when these were produced by stimulus-driven action or externally-generated events. Given prior research implicating the N2b in error monitoring processes, this result appears to reflect differences in the way that unanticipated stimuli are treated when resulting from each form of action. Within the IFM framework, error signals may serve to support dynamic motor control (Miall & Wolpert, 1996) and inform one's perception of agency over sensations (Farrer & Frith, 2002). Our findings therefore reflect differences in the degree to which these functions are served for sensations resulting from each form of action.

Because the processing of sounds produced through stimulus-driven action did not exhibit effects relating to stimulus predictability, IFM-based suppression may be unable to account for the attenuation of sensations resulting from this form of action. Self-generated stimuli (i.e., those resulting from both stimulus-driven and volitional action) exhibited reduced P2 component amplitudes, compared with stimuli resulting from externally-

generated events. Given that attentional control has previously been associated with reductions in P2 amplitude, this finding may therefore provide further support to the notion that sensory attenuation is associated attentional factors. Findings from the present study suggest that this may include differences in capacity for withdrawal of attention in particular.

Finally, the present study served to provide further caution against reliance on the N1 in examination of the sensory attenuation phenomenon. Significantly, N1 component amplitudes were not found to differ between self- and externally-generated sensations when these were made predictable in time. Although predictable tone frequencies were associated with a generalised reduction in N1 component amplitudes, evidence was identified to suggest that the effect of predictability did not differ between self- and externally-generated sensations. In each case, Bayesian analysis provided substantial evidence in favour of null hypotheses. While these findings may necessitate reconsideration of the manner in which sensory attenuation and other IFM mechanisms are studied, they also provide promising new avenues for investigation of sensorimotor effects involving the P2 and N2b.

References

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

   sensation. *Nature Neuroscience*, *1*, 635. https://doi.org/10.1038/2870

Bruin, K. J., & Wijers, A. A. (2002). Inhibition, response mode, and stimulus probability: a

   comparative event-related potential study. *Clinical Neurophysiology*, *113*(7), 1172-

   1182. https://doi.org/10.1016/S1388-2457(02)00141-4

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2 ed.). Lawrence

   Earlbaum Associates.

Colrain, I. M., Di Parsia, P., & Gora, J. (2000). The impact of prestimulus EEG frequency on

   auditory evoked potentials during sleep onset. *Canadian Journal of Experimental*

   Psychology/Revue canadienne de psychologie expérimentale, *54*(4), 243.

   https://doi.org/10.1037/h0087344

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

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

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

   component process: age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732-

   744. https://doi.org/10.1016/j.clinph.2003.11.021

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified

   theory of cognitive development and aging. *Developmental Review*, *12*(1), 45-75.

   https://doi.org/10.1016/0273-2297(92)90003-K

Desimone, R. (1996). Neural Mechanisms for Visual Memory and their Role in Attention.

   Proceedings of the National Academy of Sciences of the United States of America,

   93(24), 13494-13499. https://doi.org/10.1073/pnas.93.24.13494

Farrer, C., & Frith, C. D. (2002). Experiencing Oneself vs Another Person as Being the Cause

of an Action: The Neural Correlates of the Experience of Agency. *NeuroImage*, *15*(3),

596-603. https://doi.org/10.1006/nimg.2001.1009

Ferdinand, N. K., Mecklinger, A., & Kray, J. (2008). Error and Deviance Processing in

Implicit and Explicit Sequence Learning. *Journal of Cognitive Neuroscience*, *20*(4),

629-642. https://doi.org/10.1162/jocn.2008.20046

Ferdinand, N. K., Mecklinger, A., & Opitz, B. (2015). Learning context modulates the

processing of expectancy violations. *Brain Research*, *1629*, 72-84.

https://doi.org/10.1016/j.brainres.2015.10.017

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.

https://doi.org/10.1111/j.1469-8986.2007.00602.x

Fried, I., Haggard, P., He, B. J., & Schurger, A. (2017). Volition and Action in the Human

Brain: Processes, Pathologies, and Reasons. *The Journal of Neuroscience*, *37*(45),

10842. https://doi.org/10.1523/JNEUROSCI.2584-17.2017

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive

Sciences*, *13*(7), 293-301. https://doi.org/10.1016/j.tics.2009.04.005

Friston, K., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: a free-

energy formulation. *Biological Cybernetics*, *102*(3), 227-260.

https://doi.org/10.1007/s00422-010-0364-z

García-Larrea, L., Lukaszewicz, A.-C., & Mauguiére, F. (1992). Revisiting the oddball

paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects.

Neuropsychologia, *30*(8), 723-741. https://doi.org/10.1016/0028-3932(92)90042-K

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.

    https://doi.org/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. https://doi.org/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. *Electroencephalography and Clinical*

    Neurophysiology/Evoked Potentials Section, *92*(3), 238-252.

    https://doi.org/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. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468-

    484. https://doi.org/10.1016/0013-4694(83)90135-9

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models

    of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14-23.

    https://doi.org/10.1016/j.tics.2005.11.006

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*,

    18(4), 196. https://doi.org/10.1038/nrn.2017.14

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness.

    Nature Neuroscience, *5*(4), 382-385. https://doi.org/10.1038/nn827

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

    in the absence of movement: Differentiating motor action from sense of agency.

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

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

Harrison, A. W., Hughes, G., Rudman, G., Christensen, B. K., & Whitford, T. J. (2022). Exploring the internal forward model: Action-effect prediction and attention in sensorimotor processing. *PsyArXiv*. https://doi.org/10.31234/osf.io/rba67

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

Herwig, A., Prinz, W., & Waszak, F. (2007). Two Modes of Sensorimotor Integration in Intention-Based and Stimulus-Based Actions. *Quarterly Journal of Experimental Psychology*, *60*(11), 1540-1554. https://doi.org/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. https://doi.org/10.1126/science.182.4108.177

Hommel, B. (2000). The Prepared Reflex: Automaticity and Control in Stimulus-Response Translation. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 247-273). MIT Press.

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, *139*(1), 133-151. https://doi.org/10.1037/a0028566

James, W. (1890). *The principles of psychology*. Dover Publications.

Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations

for time and pitch. *Brain and Cognition*, *69*(1), 127-137.

https://doi.org/10.1016/j.bandc.2008.06.004

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP

experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146-157.

https://doi.org/10.1111/psyp.12639

Mauchly, J. W. (1940). Significance Test for Sphericity of a Normal n-Variate Distribution.

The Annals of Mathematical Statistics, *11*(2), 204-209.

https://doi.org/10.1214/aoms/1177731915

Mertens, R., & Polich, J. (1997). P300 from a single-stimulus paradigm: passive versus active

tasks and stimulus modality. *Electroencephalography and Clinical

Neurophysiology/Evoked Potentials Section*, *104*(6), 488-497.

https://doi.org/10.1016/S0168-5597(97)00041-5

Miall, R. C., & Wolpert, D. M. (1996). Forward Models for Physiological Motor Control.

Neural Networks, *9*(8), 1265-1279. https://doi.org/10.1016/S0893-6080(96)00035-4

Miller, G. A., Gratton, G., & Yee, C. M. (1988). Generalized implementation of an eye

movement correction procedure. *Psychophysiology*, *25*(2), 241-243.

https://doi.org/10.1111/j.1469-8986.1988.tb00999.x

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction.

Consciousness and cognition, *17*(1), 136-144.

https://doi.org/10.1016/j.concog.2006.12.004

Morey, R., & Rouder, J. (2018). BayesFactor 0.9. 12-2. *Comprehensive R Archive Network*.

http://cran.rproject.org/web/packages/BayesFactor/index.html

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on

evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313-329.

https://doi.org/10.1016/0001-6918(78)90006-9

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic

Response to Sound: A Review and an Analysis of the Component Structure.

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

8986.1987.tb00311.x

Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old

theory. *Cognition and motor processes*, 255-293. https://doi.org/10.1007/978-3-642-

69382-3_17

Novak, G., Ritter, W., & Vaughan Jr, H. G. (1992). Mismatch Detection and the Latency of

Temporal Judgments. *Psychophysiology*, *29*(4), 398-411.

https://doi.org/10.1111/j.1469-8986.1992.tb01713.x

Pfefferbaum, A., Sullivan, E. V., Rosenbloom, M. J., Mathalon, D. H., & Lim, K. O. (1998).

A Controlled Study of Cortical Gray Matter and Ventricular Changes in Alcoholic

Men Over a 5-Year Interval. *Archives of General Psychiatry*, *55*(10), 905-912.

https://doi.org/10.1001/archpsyc.55.10.905

Pfister, R., Kiesel, A., & Hoffmann, J. (2011). Learning at any rate: Action–effect learning

for stimulus-based actions. *Psychological Research*, *75*(1), 61-65.

https://doi.org/10.1007/s00426-010-0288-1

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.

https://doi.org/10.1111/psyp.1201x

Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain

Responses. *Science*, *181*(4095), 175. https://doi.org/10.1126/science.181.4095.175

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

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin*, *136*(6), 943-974. https://doi.org/10.1037/a0020541

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*(6), 482-489. https://doi.org/10.1037/h0055479

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, *37*(20), 464-476. https://doi.org/10.1007/BF00622503

Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary Observations on Tickling Oneself. *Nature*, *230*(5296), 598-599. https://doi.org/10.1038/230598a0

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*(2), 227-233. https://doi.org/10.1016/S0959-4388(98)80144-X

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*(6), 209-216. https://doi.org/10.1016/S1364-6613(97)01070-X

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880-1882. https://doi.org/10.1126/science.7569931

Zouridakis, G., Simos, P. G., & Papanicolaou, A. C. (1998). Multiple Bilaterally Asymmetric Cortical Sources Account for the Auditory N1m Component. *Brain Topography*, *10*(3), 183-189. https://doi.org/10.1023/a:1022246825461

Appendix

Table A1

Descriptive Statistics by Analysis Condition

Sequence	Task	N1		P2		N2b	
		M	*SD*	*M*	*SD*	*M*	*SD*
One-tone	Uncued Self-gen.	-3.95	3.94	1.11	3.05	0.64	3.48
	Cued Self-gen.	-3.98	3.66	0.93	3.53	0.17	4.25
	Uncued Listening	-4.80	3.41	2.76	3.07	2.04	3.10
	Cued Listening	-3.82	3.37	0.39	2.90	0.74	2.77
Two-tone	Uncued Self-gen.	-4.60	3.81	0.11	3.36	-1.10	4.03
	Cued Self-gen.	-4.65	4.25	1.21	4.53	-0.14	4.73
	Uncued Listening	-5.34	3.19	2.46	2.73	2.44	3.87
	Cued Listening	-4.51	3.51	1.96	2.50	0.72	3.09

Table A2

ANOVA Analysis of Component Amplitudes

		df_{num}	df_{den}	SS_{num}	SS_{den}	F	p	η_p^2	BF_{10}
N1	Intercept	0.98	58.85	9,682.97	4,359.78	133.26	<.001	0.69	1.00
	Sequence	0.86	51.72	49.24	249.14	11.86	.001	0.17	9.18
	Task	2.94	176.56	62.69	1,291.94	2.91	.036	0.05	1.76
	Sequence x Task	2.59	155.17	0.43	620.21	0.04	.989	0.00	0.02
P2	Intercept	0.74	44.38	1,263.58	1,722.12	44.02	<.001	0.42	1.00
	Sequence	0.77	45.95	15.09	542.04	1.67	.201	0.03	0.23
	Task	2.22	133.15	316.11	2,051.31	9.25	<.001	0.13	334,906.58
	Sequence x Task	2.30	137.84	25.46	774.34	1.97	.120	0.03	0.07
N2b	Intercept	0.82	49.11	231.76	2,488.19	5.59	0.021	0.09	1.00
	Sequence	0.80	47.73	21.04	589.62	2.14	0.149	0.03	0.26
	Task	2.46	147.33	453.50	2,547.14	10.68	0.000	0.15	4,336,194.87
	Sequence x Task	2.39	143.18	79.37	1,010.71	4.71	0.003	0.07	0.42

Table A3

Planned contrasts and interactions

		Est.	SE	95% CI Lower	Upper	df	t	p	d	BF$_{10}$
N1	Volition	0.06	0.09	-0.11	0.23	180	0.72	.472	0.11	0.19
	Self-gen.	0.11	0.12	-0.13	0.35	180	0.90	.369	0.12	0.25
	Temp. Predict	0.49	0.21	0.08	0.90	180	2.34	.020	0.30	33.62
	Volition x Sequence	-0.00	0.10	-0.20	0.19	240	-0.05	.964	-0.01	0.16
	Self-gen. x Sequence	-0.02	0.14	-0.29	0.26	240	-0.13	.896	-0.02	0.17
	Temp. Predict x Sequence	-0.08	0.24	-0.55	0.40	240	-0.32	.751	-0.04	0.11
P2	Volition	-0.22	0.11	-0.44	-0.01	180	-2.09	.038	-0.31	1,223.77
	Self-gen.	-0.54	0.15	-0.84	-0.24	180	-3.57	<.001	-0.46	1,142.17
	Temp. Predict	-0.21	0.26	-0.72	0.31	180	-0.78	.435	-0.10	0.58
	Volition x Sequence	-0.22	0.12	-0.46	0.02	240	-1.78	.077	-0.27	0.40
	Self-gen. x Sequence	0.21	0.17	-0.13	0.55	240	1.21	.229	0.16	0.26
	Temp. Predict x Sequence	-0.05	0.30	-0.63	0.54	240	-0.16	.874	-0.02	0.19
N2b	Volition	-0.09	0.12	-0.32	0.15	180	-0.72	0.475	-0.11	51.43

Self-gen.	-0.41	0.17	-0.74	-0.08	180	-2.41	0.017	-0.31	1,254.35
Temp. Predict	-0.65	0.29	-1.22	-0.08	180	-2.23	0.027	-0.29	6,692.37
Volition x Sequence	-0.44	0.13	-0.71	-0.18	240	-3.28	0.001	-0.49	3.57
Self-gen. x Sequence	-0.17	0.19	-0.54	0.20	240	-0.88	0.379	-0.11	0.21
Temp. Predict x Sequence	-0.21	0.33	-0.86	0.44	240	-0.63	0.528	-0.08	0.24

Table A4

ITI and Trial Count by Experimental Condition

Auditory Stimuli	Event Activation	Visual Stim.	ITI			Trial Count		
			M	*SD*	*Min.*	*M*	*SD*	*Min.*
Motor/Visual	Self-gen.	Cued	2,933.40	375.15	2,128.74	68.49	2.00	60
	Self-gen.	Uncued	2,916.04	356.39	2,120.75	66.89	3.49	50
	External	Cued	2,897.88	346.76	2,131.30	67.08	5.26	38
	External	Uncued	2,942.91	310.82	2,163.67	67.02	4.56	40
One-tone	Self-gen.	Cued	2,910.41	313.36	2,120.76	68.05	3.38	49
	Self-gen.	Uncued	2,922.47	340.13	2,139.94	67.34	3.70	50
	External	Cued	2,944.58	308.71	2,120.75	67.61	3.89	49
	External	Uncued	2,947.29	342.59	2,120.78	67.64	4.57	40
Two-tone	Self-gen.	Cued	2,955.33	354.33	2,125.44	68.39	2.44	57
	Self-gen.	Uncued	2,988.76	324.24	2,280.41	67.66	4.96	39
	External	Cued	2,934.07	323.45	2,126.85	68.16	3.25	48
	External	Uncued	2,918.98	315.04	2,174.28	67.74	3.85	48

Table A5

ANOVA Analysis of Inter-trial Intervals

	df_{num}	df_{den}	SS_{num}	SS_{den}	F	P	η_p^2	BF_{10}
Intercept	0.99	59.69	6.30E+09	6.29E+07	6,016.38	<.001	0.99	1.00
Auditory Stimuli	1.96	117.60	5.17E+03	4.28E+06	0.07	.930	0.00	0.02
Event Activation	0.84	50.70	1.35E+04	2.28E+06	0.35	.554	0.01	0.10
Visual Stimuli	0.94	56.39	8.13E+04	2.70E+06	1.81	.184	0.03	0.32
Auditory Stimuli x Event Activation	1.99	119.37	1.05E+04	2.93E+06	0.21	.807	0.00	0.04
Auditory Stimuli x Visual Stimuli	1.96	117.60	1.01E+03	2.64E+06	0.02	.977	0.00	0.03
Event Activation x Visual Stimuli	0.84	50.70	3.46E+04	1.02E+06	2.03	.159	0.03	0.21
Auditory Stimuli x Event Activation x Visual Stimuli	1.88	112.79	4.49E+04	3.57E+06	0.75	.472	0.01	0.11