Original Articles

Attenuation of visual evoked responses to hand and saccade-initiated flashes

Nathan G. Mifsud[a,*], Tom Beesley[a], Tamara L. Watson[b], Ruth B. Elijah[a], Tegan S. Sharp[a], Thomas J. Whitford[a]

[a] School of Psychology, UNSW Sydney, Sydney, New South Wales, Australia
[b] School of Social Sciences and Psychology, Western Sydney University, Bankstown, New South Wales, Australia

ARTICLE INFO

ABSTRACT

Sensory attenuation refers to reduced brain responses to self-initiated sensations relative to those produced by the external world. It is a low-level process that may be linked to higher-level cognitive tasks such as reality monitoring. The phenomenon is often explained by prediction error mechanisms of universal applicability to sensory modality; however, it is most widely reported for auditory stimuli resulting from self-initiated hand movements. The present series of event-related potential (ERP) experiments explored the generalizability of sensory attenuation to the visual domain by exposing participants to flashes initiated by either their own button press or volitional saccade and comparing these conditions to identical, computer-initiated stimuli. The key results showed that the largest reduction of anterior visual N1 amplitude occurred for saccade-initiated flashes, while button press-initiated flashes evoked an intermediary response between the saccade-initiated and externally initiated conditions. This indicates that sensory attenuation occurs for visual stimuli and suggests that the degree of electrophysiological attenuation may relate to the causal likelihood of pairings between the type of motor action and the modality of its sensory response.

1. Introduction

Sensory attenuation refers to self-initiated stimuli evoking reduced neurophysiological (e.g., Baess, Jacobsen, & Schröger, 2008; Houde, Nagarajan, Sekihara, & Merzenich, 2002; Schafer & Marcus, 1973) and phenomenological (e.g., Blakemore, Frith, & Wolpert, 1999; Cardoso-Leite, Mamassian, Schütz-Bosbach, & Waszak, 2010; Sato, 2008) sensory representations, compared to the sensory representations evoked by physically identical, externally initiated stimuli. The attenuation taking place here is thus related to a new external stimulus that is a consequence of enacting a motor action; a phenomenon considered here as separate to the suppression of the sensory consequences of enacting a movement within a constant sensory environment (e.g., motion across the retina produced by an eye movement).

The phenomenon is typically explained using a forward model that predicts the sensory consequences of intended actions based on internal motor commands, where these predictions are subtracted from actual sensory input (Bays & Wolpert, 2007; Wolpert, Ghahramani, & Jordan, 1995). Conversely, externally initiated stimuli lack accompanying motor information, and are thus marked by a larger disparity between predicted and actual sensory inputs—a distinction that may play a central role in cognition; specifically, our sense of agency (Engbert, Wohlschlager, & Haggard, 2008; Subramaniam, Kothare, Mizuiri, Nagarajan, & Houde, 2018). Notably, Feinberg (1978) first suggested that disruption of this distinction between self and the external world could account for some of the characteristic symptoms of schizophrenia (e.g., delusions of control), and evidence has emerged to support this theory (Ford et al., 2001; Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016; Whitford et al., 2011).

Studies of sensory attenuation have thus far largely been limited to the auditory domain (e.g., see Table 3 in Hughes, Desantis, & Waszak, 2013), centerd on a reliable event-related potential (ERP) component that is used by multiple research groups as an index of sensory attenuation—that is, the N1 or N1m component, an evoked potential or magnetic field that is consistently reduced for self-initiated vocalizations and tones (e.g., Baess et al., 2008; Curio, Neuloh, Numminen, Jousmaki, & Hari, 2000; Houde et al., 2002; Mifsud & Whitford, 2017; Sowman, Kuusik, & Johnson, 2012). Given the well-established positive relationship between the auditory N1 component and stimulus intensity (Näätänen & Picton, 1987)—i.e., loud sounds evoke larger auditory N1 amplitudes than do soft sounds—the finding that self-initiated sounds have a reduced auditory N1 response suggests that the brain processes them as being "softer". This reduced perceived loudness of self-initiated sounds may reflect an ecological adaptation, in the sense that the strong auditory feedback associated with our own speech may require attenuation to preserve the sensitivity of receptors to incoming sounds

(Bendixen, SanMiguel, & Schröger, 2012).

In contrast to the auditory domain, studies of sensory attenuation in the visual domain are relatively scarce, and results are less easily reconciled. Reported differences in visual-evoked potentials (VEPs) between self- and externally initiated visual stimuli are inconsistent in terms of both their direction and spatial location. This may be due to a diverse range of stimuli, and, relatedly, the chosen event-related components and reference electrode sites. Self-initiation has been shown to result in anterior (but not occipital) reduction of N1 for flashes (Schafer & Marcus, 1973, mastoid-referenced data) and arrow shapes (Gentsch & Schütz-Bosbach, 2011, average-referenced data), and occipital reduction of P2 for faces and houses (Hughes & Waszak, 2014, FCz-referenced data). Conversely, occipital amplification of P1 has been shown for pattern-onset stimuli (Hughes & Waszak, 2011, vertex-referenced data) and occipital amplification of N145 for pattern-reversal stimuli (Mifsud, Oestreich, et al., 2016, Fz-referenced data). However, sensory attenuation has also been observed in behavioural tasks using Gabor patches (Cardoso-Leite et al., 2010; Stenner, Bauer, Haggard, Heinze, & Dolan, 2014). The clear differences in the reported results means that further experimental work is required in the visual domain that builds on existing self-initiation paradigms.

A further limiting factor of previous studies of sensory attenuation in the visual domain is that nearly all self-initiation conditions have involved button pressing to receive a visual stimulus. This highly specific experimental condition, while relevant, must be considered in conjunction with other action–sensation contingencies to account for the wide range of circumstances that may involve sensory attenuation. In other words, there is insufficient support for assuming that findings from button-press studies can be generalized to other action–sensation contingencies. In the auditory domain, the limits of this assumption have been tested by van Elk, Salomon, Kannape, and Blanke (2014) and Mifsud, Beesley, Watson, and Whitford (2016), who employed paradigms using foot and saccade initiation respectively to demonstrate that differences in auditory-evoked potentials were dependent on the region of motor output used to produce the incoming stimulus. In the study by Mifsud, Beesley, et al. (2016), for example, a greater degree of auditory N1 attenuation was observed for button press-initiated tones than for saccade-initiated tones, consistent with the fact that while hand movements are strongly associated with auditory sensations (e.g., the sound of one's fingers typing on a keyboard), eye movements are rarely, if ever, accompanied by auditory feedback.

The present series of experiments explored whether a similar pattern of ERP effects would be observed for self-initiated visual stimuli. In Experiment 1, we tested the luminance-dependence of the frontocentral visual N1 component, and thus its conceptual similarity to the loudness-dependent auditory N1 measure used in previous sensory attenuation studies. In Experiment 2, we tested the influence of self-initiation of visual stimuli on visual N1 amplitude, extending the saccade initiation paradigm (Mifsud, Beesley, et al., 2016) to the visual domain. Lastly, in Experiment 3, we replicated the self-initiation manipulation of Experiment 2 with an added condition designed to probe the effect of temporal predictability.

2. Experiment 1

An underlying premise of neurophysiological sensory attenuation is that reductions in the ERP reflected alterations in the perceived intensity of a stimulus. For example, decreased auditory N1 amplitude has been consistently observed for self-initiated auditory tones (e.g., Sowman et al., 2012) and auditory N1 amplitude is known to decrease with decreasing stimulus intensity (Näätänen & Picton, 1987). As mentioned earlier, this suggests that self-initiated sounds are processed as "softer" (Weiss, Herwig, & Schütz-Bosbach, 2011). If we are to measure sensory attenuation in the visual domain, we must use a VEP component that reflects the intensity of a visual stimulus, in the same manner that the N1 component of the auditory-evoked potential is

sensitive to loudness. The visual N1 component is a likely candidate, as an early, sensory-evoked component that appears to be partially generated from occipital cortex (Clark, Fan, & Hillyard, 1995), but (to our knowledge) no previous studies explicitly demonstrate that the visual N1 (i.e., frontocentral maxima, mastoid-referenced) is sensitive to changes in luminance. Hence, the primary aim of Experiment 1—which only measured ERP responses to passively viewed stimuli—was to demonstrate the luminance-dependence of the visual N1.

2.1. Method

2.1.1. Participants

Eleven participants were recruited at UNSW Sydney. Six were female, 8 were right-handed, and mean age was 19 years ($SD = 1$). Participants gave written, informed consent, and received course credit in exchange for their time. This experiment, and the two that follow, were approved by the UNSW Human Research Ethics Advisory Panel (Psychology).

2.1.2. Procedure

Following provision of their demographical information, participants were fitted with an EEG cap and electrodes. EEG was continuously recorded while participants completed the experiment, seated 60 cm from a computer monitor with an integrated eye tracking system (Tobii TX300: 300 Hz gaze sampling rate; 23″, 60 Hz, 1920 × 1080 resolution TFT screen; accuracy of 0.4° visual angle; system latency under 10 ms). The eye tracking function was not used in Experiment 1, but was required for saccade detection in Experiment 2.

The experiment comprised of a series of stimulus presentations of four different types: an unstructured full-field white flash of 33.33 ms duration (i.e., two frames, verified with a photometer), that was one of two mean luminance levels, dim (10 cd/m^2) or bright (100 cd/m^2); and two types of pure tones (the data for which are not presented). The mean luminance levels were approximated based on measurements with a handheld instrument (Minolta Chroma Meter CS-100A). Participant input was not required at any time.

Each trial type was shown for 120 trials in total, intermixed in a total 480-trial sequence whose order was randomized between participants and split into 10 equal blocks separated by 30-s rest periods. Individual trials were separated by a uniformly distributed random interval (1–4 s). The EEG recording lasted approximately 30 min. Stimulus presentation was controlled by MATLAB (MathWorks, Natick, US) using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997).

EEG was recorded with a BioSemi ActiveTwo system using 64 Ag-AgCl active electrodes placed according to the extended 10–20 system. Analog signals were anti-aliased with a fixed first-order filter (−3 dB at 3600 Hz) and continuously digitized at a sampling frequency of 2048 Hz, with common mode sense (CMS) and driven right leg (DRL) used as reference and ground electrodes. During offline preprocessing, data were re-referenced to the averaged mastoid electrodes as is typical for the visual N1 (Clark et al., 1995; Vogel & Luck, 2000), band-pass filtered from 0.01 to 30 Hz (8th order zero-phase Butterworth IIR), and separated into 600-ms epochs (100 ms pre-onset and 500 ms post-onset). Data were baseline corrected with the average voltage between −100 and 0 ms. To address eye blinks and movement artefacts, we rejected individual epochs at any electrode site that contained EEG activity exceeding ± 75 µV or min-max changes in excess of 75 µV between adjacent 100-ms intervals. Individual trials were then averaged for each condition to produce ERPs for each participant. Data preprocessing was done in BrainVision Analyzer 2 (Brain Products GmbH, Munich, Germany), and statistical analyses were performed in SPSS version 23 (IBM Corp, Armonk, US).

As the latency and amplitude of flash ERPs are sensitive to stimulus parameters, a collapsed localizer approach was used to guide our analysis (Luck & Kappenman, 2012). The waveforms of the dim and bright