Title

A Crossmodal Crossover: Opposite Effects of Visual and Auditory Perceptual Load on Steady-State Evoked Potentials to Irrelevant Visual Stimuli

Author Names and Affiliations

Oscar Jacoby\textsuperscript{a,b}, Email: o.jacoby@uq.edu.au

Sarah E. Hall\textsuperscript{a,b}, Email: sarah.hall@uwa.edu.au

Jason B. Mattingley\textsuperscript{a,b}, Email: j.mattingley@uq.edu.au

\textsuperscript{a}Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia

\textsuperscript{b}School of Psychology, The University of Queensland, Brisbane, Queensland, Australia

Corresponding Author

Oscar Jacoby

E-mail: o.jacoby@uq.edu.au

Postal Address: The Queensland Brain Institute, QBI Building (#79), St Lucia, QLD 4072, Australia

Tel: +617 3346 6326

Fax: +617 3346 6301
Abstract
Mechanisms of attention are required to prioritise goal-relevant sensory events under conditions of stimulus competition. According to the perceptual load model of attention, the extent to which task-irrelevant inputs are processed is determined by the relative demands of discriminating the target: the more perceptually demanding the target task, the less unattended stimuli will be processed. Although much evidence supports the perceptual load model for competing stimuli within a single sensory modality, the effects of perceptual load in one modality on distractor processing in another is less clear. Here we used steady-state evoked potentials (SSEPs) to measure neural responses to irrelevant visual checkerboard stimuli while participants performed either a visual or auditory task that varied in perceptual load. Consistent with perceptual load theory, increasing visual task load suppressed SSEPs to the ignored visual checkerboards. In contrast, increasing auditory task load enhanced SSEPs to the ignored visual checkerboards. This enhanced neural response to irrelevant visual stimuli under auditory load suggests that exhausting capacity within one modality selectively compromises inhibitory processes required for filtering stimuli in another.

Keywords
Distractor interference, frequency tagging, neural oscillations, perceptual load, selective attention
1. Introduction

The factors that influence the extent to which we are able to selectively process only attended information and suppress unattended or task-irrelevant information has been the subject of much interest and controversy in cognitive neuroscience (Broadbent, 1958; Deutsch and Deutsch, 1963; Driver, 2001; Lachter et al., 2004; Pashler, 1998; Yantis, 2000). According to Lavie and colleagues’ perceptual load theory (Lavie, 1995; Lavie and Tsal, 1994), one factor that influences the extent to which unattended stimuli are processed is the load placed on the perceptual system by the task at hand. Perceptual load theory assumes that perceptual resources have a limited capacity, and also that they are involuntarily applied to as much of the environment as possible. Thus, if a task currently does not absorb all perceptual resources, any remaining capacity is automatically allocated to the processing of other, potentially irrelevant or distracting, stimuli. As task load increases, fewer perceptual resources are available to be allocated to other stimuli. Here we used frequency tagged stimuli and steady-state evoked potentials (SSEPs) to measure the effects of visual and auditory perceptual load on neural responses to competing, irrelevant stimuli in the peripheral visual field.

A large body of evidence supports the predictions of perceptual load theory in situations in which both attended and unattended stimuli are presented exclusively within the visual modality (Lavie, 2005, 2010; Lavie et al., 2004). For instance, numerous behavioural studies have examined the relationship between visual perceptual load and the magnitude of the flanker compatibility effect (Eriksen and Eriksen, 1974). These studies have shown that as the perceptual load associated with the target is increased, the magnitude of the flanker compatibility effect is reduced or even eliminated altogether (Lavie, 1995; Lavie and Cox, 1997; Lavie et al., 2004; Lavie et al., 2003), consistent with the notion that increasing perceptual load reduces involuntary processing of unattended flanker items. Further evidence
Crossmodal perceptual load

comes from studies using functional magnetic resonance imaging (fMRI) to measure the magnitude of blood-oxygen-level dependent (BOLD) responses to unattended visual stimuli under varying levels of visual perceptual load. These studies have demonstrated that increasing perceptual load reduces BOLD responses to unattended visual stimuli across a range of brain areas along the visual processing pathway, including the lateral geniculate nucleus (O’Connor et al., 2002), areas V1, V2, V3, and V4 (Schwartz et al., 2005), area TEO (Pinsk et al., 2004), and area V5 (Rees et al., 1997).

In contrast to situations in which both attended and unattended stimuli occur in the same modality, the effect of perceptual load on the processing of unattended stimuli in different modalities has not been widely explored, and the findings of the few studies that have examined the issue are mixed (Driver and Spence, 1998; Spence, 2010). Klemen, Büchel, and Rose (2009) examined the effect of auditory perceptual load on neural responses to task-irrelevant visual stimuli using fMRI. Participants performed a one-back matching task on serially presented streams of auditory tones while passively viewing task-irrelevant images of natural scenes. The perceptual load of the auditory matching task was manipulated by varying the number of different-frequency tones in each stream, and the frequency difference between each tone. BOLD responses associated with the unattended visual images were reduced in lateral occipital cortex (LOC) under conditions of high- relative to low-auditory perceptual load. This finding suggests that perceptual load effects arise not only for exclusively visual targets and distractors, but also for competing stimuli across auditory and visual modalities, consistent with theoretical models of crossmodal attention that argue that perceptual capacity limits arise at a supramodal level (Broadbent, 1958; Kahneman, 1973).

Further evidence supporting supramodal theories of attention comes from studies that have examined the effect of auditory search tasks on the processing of irrelevant visual motion (Berman and Colby, 2002; Houghton et al., 2003). For example, when participants
perform a perceptually demanding search task on either visual or auditory search items, BOLD responses are reduced to an irrelevant visual motion stimulus in motion-sensitive area MT+, as is the subjective duration of the motion aftereffect (Mather et al., 2008) induced by the motion stimulus.

In contrast to these crossmodal findings, other studies have failed to find evidence that task-related perceptual load in one modality modulates neural responses to unattended stimuli in another. Rees, Frith, and Lavie (2001) examined the effect of auditory perceptual load on neural responses to an irrelevant visual motion stimulus. Participants monitored auditory streams of spoken words for either loudly spoken words within softly spoken words (low perceptual load), or bisyllabic words within monosyllabic and trisyllabic words (high perceptual load). Neural responses in area V5/MT to a concurrently presented visual motion stimulus were measured using positron emission tomography (PET), and found to be similar in magnitude between auditory perceptual load conditions. This finding runs counter to supramodal accounts of perceptual load effects, but fits well with theoretical models arguing for modality specificity of attentional resources (Allport et al., 1972; Duncan et al., 1997; Mcleod, 1977; Treisman and Davies, 1973; Wickens, 1980).

Further evidence for such modality-specific accounts was recently provided by Parks, Hilimire, and Corballis (2011), who examined the issue of crossmodal perceptual load effects using steady-state evoked potentials (SSEPs). SSEPs are cyclical oscillations of activity in sensory cortical areas evoked by a repeatedly presented or flickering stimulus stream (Di Russo et al., 2007; Heinrich, 2010; Picton et al., 2003; Toffanin et al., 2009; Vialatte et al., 2010). Such oscillations are evoked by both visual and auditory stimulation (e.g., Keitel et al., 2011; Saupe et al., 2009). The amplitude of these oscillations can be mathematically extracted from the scalp-recorded electroencephalogram (EEG) using a fast Fourier transform (FFT). If multiple stimulus streams are presented simultaneously, each at a unique frequency,
the SSEP amplitude observed at each frequency can be used to measure the strength of the neural response to each stimulus stream independently over an extended period of time (Toffanin et al., 2009). Parks et al. (2011) took advantage of this feature of the SSEP technique to simultaneously measure neural responses to both an attended stream of coloured crosses and an irrelevant stimulus presented in either the visual or auditory modality. Perceptual load was manipulated by having participants search for targets defined either on the basis of a single feature (colour), or on the basis of a conjunction of two features (colour and orientation). Consistent with perceptual load theory (Lavie, 1995), increasing visual load increased SSEP responses to the attended stream, and reduced SSEP responses to unattended peripheral checkerboard stimuli. In contrast, and consistent with modality-specific accounts of attention (Allport et al., 1972; Duncan et al., 1997; Mcleod, 1977; Treisman and Davies, 1973; Wickens, 1980), increasing visual load had no effect on SSEP responses to unattended auditory stimuli.

Further investigations are clearly necessary to conclusively establish the relationship between perceptual load and crossmodal distractor processing. Methodological differences in the studies conducted thus far further complicate interpretation. One factor which may influence the nature of the relationship is the modality of the load task. It is possible, for example, that the influence of visual perceptual load on auditory distractors qualitatively differs from the influence of auditory perceptual load on visual distractors. Caution must also be exercised in accepting null results, such as those observed by Rees et al. (2001) and Parks et al. (2011).

In the present study, we set out to investigate the effects of visual and auditory perceptual load on the processing of competing visual stimuli. In two separate experiments, participants performed search tasks of varying perceptual load on serially presented streams of either visual or auditory search items, while a task-irrelevant checkerboard flickered in the
peripheral visual field. Neural responses to both attended and unattended stimuli were measured using SSEPs, and compared across levels of perceptual load. In contrast to the study by Parks et al. (2011), we employed the same unattended (visual) stimuli in both experiments to permit a direct comparison of the effects of visual and auditory load on neural responses to a constant unattended stimulus. In line with perceptual load theory (Lavie, 1995), we expected that increases in visual perceptual load would give rise to reduced SSEP responses to the unattended visual checkerboard, as fewer residual perceptual resources would be available to process them. We also expected that increasing visual perceptual load would enhance SSEP responses evoked by the attended stimulus stream, in line with evidence that increasing the perceptual load imposed by an attended stimulus enhances neural responses to that stimulus (Parks et al., 2011; Pinsk et al., 2004).

If perceptual capacity limits arise at a supramodal level (Berman and Colby, 2002; Broadbent, 1958; Houghton et al., 2003; Kahneman, 1973; Klemen et al., 2009; Parks et al., 2009) and perceptual load theory applies crossmodally as well as within a modality, increasing auditory perceptual load should reduce SSEPs to the unattended visual checkerboard, just as predicted for visual perceptual load. Alternatively, if capacity-limitations underlying the perceptual load effect are entirely modality-specific (Allport et al., 1972; Duncan et al., 1997; Mcleod, 1977; Parks et al., 2011; Rees et al., 2001; Treisman and Davies, 1973; Wickens, 1980), then increasing auditory perceptual load should have no effect on SSEP responses to the unattended visual checkerboard. As with the visual load experiment, we also predicted that increasing the perceptual load associated with the attended (auditory) stimulus stream would enhance SSEP responses evoked by these stimuli.
2. Materials and Methods

2.1. Participants

Twenty-five participants (16 female, age $M = 21.6$ years, $SD = 2.2$ years) took part in the within-modality (visual load) experiment, and 26 participants (16 female, age $M = 22.7$ years, $SD = 3.7$ years) took part in the between-modality (auditory load) experiment. One individual participated in both experiments. All participants were right-handed and had normal or corrected-to-normal vision. All experimental procedures were conducted in accordance with the principles expressed in the Declaration of Helsinki, and were approved by the University of Queensland Ethics Committee. Written informed consent was obtained from each participant prior to participation. Participants were financially reimbursed at a rate of AU$10 per hour.

2.2. Stimuli and Procedure

Visual stimuli were presented on a 21-inch CRT monitor (NEC, Accusync 120) at a screen resolution of $1024 \times 768$ pixels and a refresh rate of 100 Hz. All visual stimuli were superimposed on a mid-grey background (RGB coordinates 127, 127, 127). Auditory stimuli were presented via a pair of computer speakers located to the left and right of the computer monitor. Participants were seated at a viewing distance of 57 cm from the monitor and speakers. Stimulus presentation and response recording were controlled using Cogent software (Cogent 2000 toolbox: FIL, ICN, and Wellcome Department of Imaging Neuroscience) in Matlab version 7.8 (www.mathworks.com), running on a Pentium IV 3GHz desktop computer. Trials in each experiment involved the presentation of two flickering streams of stimuli for 6 s; an attended stream and an unattended stream. Each stream of stimuli is described in detail below.

**Unattended visual checkerboard.** In both experiments, the unattended stimulus consisted of two counterphasing “wedges” of a checkerboard, one presented in each visual
Crossmodal perceptual load

Figure 1 contains examples of the attended and unattended stimuli used in the visual load experiment. Each wedge spanned an area ranging between 2.5° and 13.8° of visual angle from fixation, and between -72° and 72° of polar angle from the horizontal meridian. These wedges were divided into checkerboard segments using concentric rings and radial spokes. The spacing between concentric rings was 1.1° of visual angle, and the spacing between radial spokes was 12° of polar angle. The segments of the checkerboard stimulus counterphased between dark grey (RGB 64, 64, 64) and light grey (RGB 192, 192, 192) at a rate of 10.0 Hz. A counterphasing checkerboard stimulus was used for the unattended visual stimulus stream because previous studies have demonstrated that such stimuli evoke strong neural responses across a large number of visual areas (O’Connor et al., 2002; Parks et al., 2011; Schwartz et al., 2005).

Figure 1. Schematic representation of the stimulus streams used in the visual load experiment. (A) Participants monitored a central stream of coloured crosses alternating at 4 Hz while ignoring a peripheral checkerboard flickering at 10 Hz. (B) In the no-load condition, participants passively observed the crosses. In the low-load
condition, participants counted the red crosses. In the high-load condition, participants counted the upright yellow and inverted green crosses. The streams of crosses were identical across the three load conditions. (Note: the crosses are shown here in greyscale but were coloured in the actual task.)

**Attended visual crosses.** In the visual load experiment, participants were required to monitor a central stream of coloured crosses at fixation. Crosses appeared in one of two orientations (upright or inverted), and one of eight different colours (yellow, RGB 255, 255, 0; pink, RGB 255, 0, 255; light blue, RGB 0, 255, 255; red, RGB 255, 0, 0; green, RGB 0, 255, 0; dark blue, RGB 0, 0, 255; white, RGB 255, 255, 255; and black, RGB 0, 0, 0). The lines making up each cross were 0.3° thick, and the vertical line was 2.5° high. The horizontal line was 1.5° wide, and intersected the vertical line 0.4° above (upright crosses) or below (inverted crosses) the midline. Crosses were presented upon a circular pedestal extending out to 2.5° of visual angle from fixation. The attended visual stimulus changed at a rate of 4.0 Hz. On each change, the colour and/or orientation of the cross changed, and the colour of the circular pedestal alternated between dark grey (RGB 64, 64, 64) and light grey (RGB 192, 192, 192).

In separate blocks of trials, participants performed one of three tasks, corresponding to the no-load, low-load and high-load conditions. In the no-load condition, participants passively fixated the central stream of crosses. In the low-load condition, participants monitored the cross stream for all red crosses, irrespective of orientation (feature search). In the high-load condition, participants monitored the stream for either of two possible targets: an upright yellow cross or an inverted green cross (conjunction search; see Schwartz et al. 2005).

**Attended auditory phonemes.** In the auditory load experiment, participants were required to monitor streams of spoken phonemes (see Figure 2). The phonemes (‘Ba’, ‘Da’, ‘Ga’, ‘Koo’, and ‘Bee’) were presented in either a male or female voice. In some instances
within the auditory stream, the pitch of the phoneme was digitally distorted to be noticeably higher or lower than a normal speaking voice (by an average of 40%) using WavePad Audio Editing Software (www.nch.com.au/wavepad/index.html). As with the visual cross stimuli, the auditory phonemes were presented at a rate of 4.0 Hz. Although auditory SSEPs tend to be optimally evoked at higher frequencies of around 40 Hz (Picton et al., 2003), reliable auditory SSEPs have also been observed at lower frequencies similar to that used here (e.g., Nozaradan, Peretz, Missal et al., 2011; Nozaradan, Peretz and Mouraux, 2011). Additionally, we felt that matching the frequency of the visual and auditory tasks was more important than evoking the strongest possible auditory SSEPs. To ensure that the auditory stimuli evoked regular neural activity at the intended frequency, an audible click was presented synchronously with the onset of each phoneme. As with the visual load experiment, participants completed each of the three load conditions in separate blocks of trials. In the no-load condition, participants listened passively to the auditory stream without performing any task. In the low-load condition, participants monitored the auditory stream for any pitch-distorted phoneme (feature search). In the high-load condition, participants monitored the auditory stream for any female-spoken ‘Bee’ phoneme and any male-spoken ‘Da’ phoneme (conjunction search). These tasks constituted auditory analogues of the feature and conjunction searches implemented in the low- and high-visual load conditions, respectively.
Figure 2. Schematic representation of the stimulus streams used in the auditory load experiment. (A) Participants fixated on a static central fixation cross while an irrelevant checkerboard – identical to that used in the visual load experiment – flickered at 10 Hz in the periphery. Participants monitored an auditory stream of spoken phonemes delivered via audio speakers at a rate of 4 Hz (i.e., one phoneme every 250 ms). Positioning the speakers to the left and right of the monitor was done to give the impression of sounds coming from the centre of the monitor. (B) Phonemes could be presented in either a male voice (shown in black) or female voice (shown in white). Some of the phonemes were digitally distorted such that their pitch was noticeably higher (top row) or lower (bottom row) than a normal speaking voice (middle row). In the no-load condition, participants passively listened to the phonemes. In the low-load condition, participants counted the pitch-distorted phonemes. In the high-load condition, participants counted the male ‘Da’ and female ‘Bee’ phonemes. Note that the actual streams of phonemes were identical across the three load conditions, as for the central stream in the visual load experiment.

For both the visual- and auditory-load experiments, the number of target items that appeared within a trial was zero, one, or two, with equal probability during each block. The temporal positions of any target items within each trial were pseudorandomised, with three
constraints. First, the earliest a target could appear was 750 ms after the onset of the attended stream (i.e., the fourth item in the stream). This was done to avoid situations in which participants might miss a target because they were temporarily distracted by the onset of the flickering stimuli. Second, on trials in which there were two targets, a minimum stimulus onset asynchrony (SOA) of 750 ms was maintained between targets (i.e., a minimum of two non-targets were interposed between targets). This was done to avoid situations in which participants might miss a second target due to an “attentional blink” (where identifying an initial masked visual target tends to impair identification of a subsequent masked target presented within 500 ms of the first; Shapiro et al., 1997). Third, on trials in which there were two targets, the earliest the second target could appear was 5000 ms after the onset of the attended stream (i.e., only after the fifth-last item in the stream), so participants would be discouraged from prematurely disengaging from the visual stimuli because all possible targets had already appeared in the stream.

At the end of each trial, participants were prompted to report the number of target items they had detected by pressing buttons on a computer keyboard. In the no-load condition, participants were instructed to press any key at the end of the trial. Visual feedback was then provided via a green (for correct) or red (for incorrect) dot, which appeared centrally for 500 ms after participants made their response. The next trial began after a 500 ms blank, mid-grey screen. Each participant completed three blocks of 45 trials for each of the three load conditions, with block order counterbalanced across participants. For each participant, the trials for each load condition involved identical stimulus streams, although the order in which they were presented was randomised separately for each condition. Prior to completing the main experiment, participants completed one practice block of 18 trials for each of their three load conditions. Participants were instructed as to which targets they were looking for in each block via on-screen instructions presented at the start of each block.
2.3. EEG Recording

Continuous EEG data were recorded using a BioSemi Active Two system (BioSemi, Amsterdam, Netherlands), digitized at a 1024 Hz sample rate with 24-bit A/D conversion. The 64 active scalp Ag/AgCl electrodes were arranged according to the international standard 10-10 system for electrode placement (Oostenveld and Praamstra, 2001) using a nylon head cap. As per BioSemi system design, the Common Mode Sense and Driven Right Leg electrodes served as the ground, and all scalp electrodes were referenced to the Common Mode Sense electrode during recording. Eye movements were monitored using bipolar horizontal electrooculographic (EOG) electrodes placed at the outer canthi of each eye, and bipolar vertical EOG electrodes placed above and below the left eye.

2.4. EEG Data Analysis

Offline EEG data analysis was performed using Brain Electrical Source Acquisition (BESA 5.3; MEGIS Software GmbH, Gräfelfing, Germany). The data for the scalp electrodes were re-referenced to the average of all 64 scalp electrodes, and subjected to a 0.1 Hz high-pass digital filter. Noisy channels, identified by visual inspection of the data, were replaced by a spherical spline interpolation of the voltages recorded at all other scalp electrodes. The maximum number of electrodes interpolated in any participant was four. The data were then segmented into 6 s epochs spanning the duration of each trial. Epochs containing excessive noise created by eye blinks, swallowing, or other muscle activity were identified by visual inspection of the EEG and EOG data and excluded from further analysis. Trials in which the participant responded incorrectly were also excluded from the analysis, on the grounds that the source of error on these trials is unknown. Data from two participants in each experiment were excluded from further analysis because more than 35% of their trials had to be rejected for violating these criteria. Data from a third participant in the auditory load experiment were excluded from further analysis because the participant did not correctly follow task
instructions. For the remaining participants, an average of 20% of trials were rejected for violating the above criteria.

For each participant, the remaining epochs were then averaged together, separately for each of the three load conditions (no, low, high). Averaging the data across multiple epochs should have removed any activity in the continuous EEG data not time-locked to the visual stimuli (Luck, 2005). The data in these averaged waveforms between 750 and 4750 ms after the onset of the flickering stimuli were then subjected to an FFT using a purpose-built Matlab script. The first 750 ms after the onset of the flickering stimuli was excluded from the analysis to allow stabilisation of the oscillatory neural responses to the flickering stimuli (Morgan et al., 1996). The final 1250 ms of each trial was also excluded because this was the time during which the second target appeared in trials in which two targets were presented. Removing this final section of data allowed us to avoid inclusion of data recorded when participants were potentially no longer actively engaged in the load task.

From the FFTs, we identified the SSEP amplitudes corresponding to the flicker frequencies of the attended (4 Hz) and unattended (10 Hz) stimulus streams, separately for each of the 64 scalp electrodes. In line with previous SSEP studies (Di Russo et al., 2007; Ding et al., 2006; Parks et al., 2011), SSEP amplitudes evoked by the unattended (visual) stimulus stream tended to be largest at posterior electrodes located over the occipital lobes for visual stimuli, and at frontal electrodes for auditory stimuli (see Figure 3). On the basis of these observations, we analysed the visual stimulus SSEPs at a group of three occipital electrodes (O1, Oz, and O2), and the auditory stimulus SSEPs at a group of three frontal electrodes (F1, Fz, and F2).
Figure 3. Scalp distributions of SSEP amplitudes at the flicker frequencies of attended (4 Hz) and unattended (10 Hz) stimuli in each experiment, collapsed across the three load conditions. Colours correspond to the colour bar to the right of each head map. The purple circles indicate the scalp locations of the three electrodes used in each analysis. Images were created using a Matlab script adapted from EEGLAB (Delorme and Makeig, 2004).

3. Results

All statistical tests were conducted using SPSS, with a two-tailed alpha level of .05. Mauchly’s Test of Sphericity was applied to all within-subjects F-tests. Greenhouse-Geisser epsilon adjustments were made to degrees of freedom for these F-tests wherever the
assumption of sphericity was untenable. Unadjusted degrees of freedom are reported for all $F$-tests.

### 3.1. Behavioural Results

To evaluate the effectiveness of our load manipulations, within-subjects $t$-tests were conducted to compare accuracy (the proportion of trials in which participants correctly reported the number of targets presented) between the low- and high-load conditions, separately for each experiment. For the visual load experiment, accuracy was significantly higher in the low-load condition ($M = 96.4\%, SD = 2.6\%$) than in the high-load condition ($M = 76.4\%, SD = 11.6\%$), $t(22) = 8.79, p < .001$. Similarly, for the auditory load experiment, accuracy was significantly higher in the low-load condition ($M = 92.9\%, SD = 6.8\%$) than in the high-load condition ($M = 59.4\%, SD = 9.7\%$), $t(22) = 17.36, p < .001$. These results demonstrate that participants in both experiments found the high-load task more difficult than the low-load task, supporting the effectiveness of our load manipulations.

### 3.2. SSEP Amplitude Results

To investigate the effects of increasing perceptual load on the strength of neural representations of attended and unattended stimuli, we subjected the SSEP amplitudes of each stimulus stream to $3 \times 3$ within-subjects ANOVAs with factors of load (no, low and high) and electrode, separately for each experiment. For clarity, we report here only the results of statistical effects that have direct relevance for the experimental hypotheses.

#### 3.2.1. Visual load experiment

Figure 4A and Figure 4C represent SSEP amplitudes as a function of load for the attended (4Hz) and unattended (10Hz) stimuli, respectively. Frequency spectrum data centred on the flicker frequencies of the attended and unattended stimuli are presented in Figure 4B and Figure 4D, respectively. As shown in Figure 4C, SSEP amplitudes for the unattended checkerboards decreased with increasing load, consistent with load theory (Lavie, 1995). Statistically, a significant main effect of load was present for SSEP
amplitudes corresponding to the unattended checkerboards, \( F(2,44) = 5.35, p = .008, \eta^2 = .007 \). Follow-up within-subjects \( t \)-tests revealed that SSEPs evoked by the peripheral checkerboards were significantly larger in the no-load condition (\( M = 2.64 \mu V, SD = 1.40 \mu V \)) than in the high-load condition (\( M = 2.33 \mu V, SD = 1.33 \mu V \)), \( t(22) = 3.38, p = .003 \). SSEP amplitudes for the peripheral checkerboards did not differ significantly between the no-load condition and the low-load condition (\( M = 2.50 \mu V, SD = 1.39 \mu V \)), \( t(22) = 1.83, p = .080 \), or between the low- and high-load conditions, \( t(22) = 1.47, p = .155 \). There was no significant effect of load on SSEP amplitudes corresponding to the attended crosses, \( F(2,44) = 2.47, p = .097, \eta^2 = .009 \).

**Figure 4.** SSEP results for the visual load experiment. (A) Mean SSEP amplitudes at the flicker frequency of the attended cross stream (4 Hz) as a function of load, collapsed across electrodes O1, Oz, and O2. Error bars represent within-subjects standard error of the mean (Loftus and Masson, 1994). (B) Frequency spectrum data
between 3 and 5 Hz for each load condition, collapsed across electrodes O1, Oz, and O2. (C) Mean SSEP amplitudes at the flicker frequency of the unattended checkerboard stimulus (10 Hz) as a function of load, collapsed across electrodes O1, Oz, and O2. (D) Frequency spectrum data between 9 and 11 Hz for each load condition, collapsed across electrodes O1, Oz, and O2.

3.2.2. Auditory load experiment. Figure 5A and Figure 5C represent SSEP amplitudes as a function of load for the attended (4 Hz) and unattended (10 Hz) stimuli, respectively. Frequency spectrum data centred on the flicker frequencies of the attended and unattended stimuli are presented in Figure 5B and Figure 5D, respectively. As shown in Figure 5C, SSEP amplitudes for the unattended stimuli increased with increasing load. Statistically, there was a marginally significant main effect of load on SSEP amplitudes corresponding to the peripheral checkerboards, $F(2,44) = 3.39, \epsilon = .75, p = .058, \eta^2 = .001$. Follow-up within-subjects $t$-tests revealed that SSEP amplitudes evoked by the peripheral checkerboards were significantly smaller in the low-load condition ($M = 2.41 \mu V, SD = 1.46 \mu V$) than in the high-load condition ($M = 2.53 \mu V, SD = 1.56 \mu V$), $t(22) = 2.14, p = .044$. SSEP amplitudes for the peripheral checkerboards in the no-load condition ($M = 2.39 \mu V, SD = 1.43 \mu V$) were also marginally smaller than those evoked in the high-load condition, $t(22) = 1.96, p = .063$, but did not differ significantly from those evoked in the low-load condition, $t(22) = 0.52, p = .609$. There was no significant effect of load on SSEP amplitudes corresponding to the attended phonemes, $F(2,44) = 0.11, \epsilon = .66, p = .807, \eta^2 = .001$. 

Crossmodal perceptual load
Figure 5. SSEP results for the auditory load experiment. (A) Mean SSEP amplitudes at the flicker frequency of the attended phoneme stream (4 Hz) as a function of load, collapsed across electrodes F1, Fz, and F2. Error bars represent within-subjects standard error of the mean. (B) Frequency spectrum data between 3 and 5 Hz for each load condition, collapsed across electrodes F1, Fz, and F2. (C) Mean SSEP amplitudes at the flicker frequency of the unattended checkerboard stimulus (10 Hz) as a function of load, collapsed across electrodes O1, Oz, and O2. (D) Frequency spectrum data between 9 and 11 Hz for each load condition, collapsed across electrodes O1, Oz, and O2.

4. Discussion

We used SSEPs to measure neural responses to irrelevant visual checkerboard stimuli while participants performed either visual or auditory tasks that varied in perceptual load. We found that increasing visual task load decreased the magnitude of SSEP responses to the ignored visual checkerboards, but had no effect on SSEP responses to the attended visual
stimuli. These findings are broadly consistent with the assumptions of load theory (Lavie, 1995), according to which perceptual resources not allocated to target processing spill over to distractor elements until capacity is exhausted. In the present context, as the perceptual demands of the visual task increased, fewer resources were available to process the unattended checkerboard.

A strikingly different pattern emerged when the central task involved auditory stimuli. Here, rather than attenuating neural responses to the irrelevant checkerboard, SSEPs were actually enhanced for the visual stimuli with increases in auditory load. These distinct and opposite effects of visual and auditory task load on irrelevant visual stimulus processing cannot be explained by physical stimulus differences, as the same checkerboards were used in each experiment. The presentation rate of the irrelevant checkerboards was also matched across the two experiments, ruling out the possibility that the different effects were due to different cortical networks being selected by the tagged flicker frequencies (Ding et al., 2006). Significant reductions in accuracy in the high- relative to low-load conditions of each experiment indicate that both task manipulations were effective at depleting resources. Response demands were also matched: participants had to detect zero, one or two targets per trial an equal number of times in each experiment. Our finding of increased processing of irrelevant visual stimuli with increasing auditory task load cannot be satisfactorily explained by either supramodal (e.g., Broadbent, 1958; Kahneman, 1973) or modality-specific (e.g., Allport et al., 1972; Wickens, 1980) accounts of attention.

The increase in SSEP responses to the unattended checkerboards in the auditory task experiment is also unlikely to result from a generalised increase in arousal with increasing auditory task load. First, there is evidence that increasing arousal decreases rather than increases steady-state evoked responses to auditory stimuli (Griskova et al., 2007). Second, one would expect a generalised increase in arousal to influence responses to both attended
and unattended stimuli, whereas in the present experiment SSEP responses to the attended phonemes themselves were not affected by auditory task load.

A similar positive relationship between perceptual load in one modality and distractor processing in another was reported by Tellinghuisen and Nowak (2003). They found that increasing visual perceptual load eliminated the flanker compatibility effect evoked by irrelevant visual flanker letters, but enhanced the flanker compatibility effects evoked by auditory distractors. Tellinghuisen and Nowak explained their findings by proposing that some attentional resources involved in prioritising the processing of task-relevant stimuli are also involved in inhibiting the processing of potentially distracting information from currently unattended modalities. According to this account, increasing task-related perceptual load in one modality has two distinct effects on the processing of unattended stimuli. First, as stated by perceptual load theory, it should reduce the amount of residual (modality-specific) perceptual capacity available to process other stimuli in the same modality, thereby decreasing unattended stimulus processing in that modality. Second, it should impinge on the attentional resources required to inhibit stimulus processing in other modalities, thereby increasing the processing of stimuli present in those modalities. This crossmodal disinhibition account fits well with theoretical models arguing that a key role of attentional resources is the suppression of task-irrelevant stimulus processing (e.g., Desimone and Duncan, 1995; Pashler, 1998). On this account, neural responses (and therefore SSEPs) to the irrelevant visual checkerboards would effectively be released from inhibition as the load of the central auditory task was increased, as resources normally involved in inhibiting irrelevant stimuli were re-allocated to processing the task-relevant auditory stimuli.

Our findings go beyond the behavioural results of Tellinghuisen and Nowak (2003) by providing a direct neural marker for the crossmodal disinhibition effect, and permit us to rule out a number of caveats to the generality of their theory. First, whereas Tellinghuisen
and Nowak measured the effect of visual perceptual load on auditory distractor processing, we measured the effect of auditory perceptual load on visual distractor processing. This allowed us to demonstrate that the effects occur when the load task is visual or auditory, and also to directly compare the crossmodal effects to those observed in a unimodal (visual) situation using identical distractor stimuli. Second, unlike Tellinghuisen and Nowak, who used a flanker task in which both target and distractor stimuli were potentially task-relevant (both stimulus types could be mapped to speeded behavioural responses), we used distractor stimuli that had no meaningful association with participants’ task, and could never be confused with elements within the auditory stream.

The selective enhancement of crossmodal distractor processing with increasing perceptual load observed here and by Tellinghuisen and Nowak (2003) stands in contrast to other studies in which increasing perceptual load in one modality either reduced distractor processing in another modality (Berman and Colby, 2002; Houghton et al., 2003; Klemen et al., 2009), or had no measurable effect (Parks et al., 2011; Rees et al., 2001). One factor that might have contributed to these different patterns of results is the nature of the load manipulations employed in each study. Evidence from unimodal load protocols has revealed that different task manipulations can have opposite effects on distractor processing: perceptually demanding tasks, such as those that require a target to be distinguished from a group of similar distractors, tend to reduce distractor processing, whereas cognitively demanding tasks that require executive control, as occurs during working memory manipulations, typically increase distractor processing (Kelley and Lavie, 2011; Lavie, 2005, 2010; Lavie et al., 2004). In the present study, we attempted to match our auditory and verbal tasks for working memory load, both in terms of the rules that defined targets and in the number of potential targets to be identified per trial (0, 1 or 2). The relative extent to which previous task manipulations taxed cognitive resources, perceptual resources, or both, might
have influenced the magnitude and direction of the observed effects on crossmodal distractor processing. Additionally, task manipulations that involve reducing the “quality” of the presented stimuli (such as reducing the frequency difference between target and non-target tones across conditions of load, e.g., Klemen et al., 2009) represent cases of signal data limitations (Norman and Bobrow, 1975) as opposed to perceptual load limitations. Care must be taken when comparing the effects of these signal data limitations with the effects of more conventional perceptual load manipulations (Lavie and Tsal, 1994).

It is also worth considering differences in the specific neural processes that are measured by brain imaging techniques such as fMRI, PET and EEG. Attentional effects can vary in magnitude and presence across brain regions (e.g., Kastner et al., 1999; Liu et al., 2005), and also across time within the same brain region (e.g., Di Russo et al., 2003; Noesselt et al., 2002). Similarly, perceptual load may not influence crossmodal distractor processing in the same way across different neural structures and processing stages. Further work using combined imaging techniques, such as concurrent EEG and fMRI, will be needed to determine how the neural signatures of load vary across distinct brain regions over time.

Finally, previous findings of no relationship between load and crossmodal distractor processing could potentially be due simply to a lack of statistical power, for example due to weak load manipulations, insufficient numbers of trials or participants, or a relative insensitivity of the recorded responses. We note in this context that in their study of audiovisual perceptual load, Parks et al. (2011; see Figure 4) observed numerical increases in SSEP responses to auditory distractors with increasing visual perceptual load at each of the four electrodes analysed, though these differences were not statistically reliable. Similarly, Rees et al. (2001; Experiment 2) found non-significant increases in the duration of the motion aftereffect induced by an irrelevant motion stimulus with increasing auditory task load in three of their four participants.
One possible alternative explanation for the findings of the present study relates to our decision to present the auditory task stimuli via speakers positioned to the left and right of the computer monitor. Although we had hoped that this would create the impression of a central sound source (comparable to the central stream of crosses in the visual task), it may have been the case that presenting auditory stimuli in the periphery caused participants to distribute their attention diffusely over both speakers and the monitor. If this were the case, SSEP responses to the unattended visual stimuli may have increased with increasing auditory load because of crossmodal links in the deployment of spatial attention (Driver & Spence, 1998; Eimer & Schröger, 1998; McDonald et al., 2003). A future study similar to the one reported here but using a single central speaker for the delivery of auditory stimuli would usefully address this possibility.

A better understanding of the crossmodal effects of perceptual load will require further research involving systematic manipulation of the factors that differ between the studies that have attempted to examine the issue so far. These include: the modality of the attended and unattended stimuli; the nature of the task load manipulation; whether unattended stimuli can be meaningfully associated with the task; the sensitivity of the stimulus processing measure; and the specific neural processes measured. Nonetheless, the findings reported here reinforce the suggestion of Tellinghuisen and Nowak (2003) that, paradoxically, devoting more perceptual resources to stimuli presented in one modality can compromise the ability to inhibit the processing of stimuli present in other modalities.
Funding

This work was supported by the Australian Research Council (grant number DP110102925 to J.B.M). The sponsors had no involvement in the study design, data collection, analysis, and interpretation, writing of the report, or decision to submit the article for publication.

Acknowledgements

We thank Dr David Lloyd and Mr David Painter for assistance in developing the stimulus presentation and EEG data analysis software used in this project. We also thank Dr Marc Kamke and Dr Martin Sale for useful comments on the manuscript.
References


responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral Cortex, 15*(6), 770-786.


