Dynamic crossmodal links revealed by steady-state responses in auditory–visual divided attention

Ritske de Jong *, Paolo Toffanin, Marten Harbers

Experimental Psychology, University of Groningen, The Netherlands

ARTICLE INFO

ABSTRACT

Frequency tagging has been often used to study intramodal attention but not intermodal attention. We used EEG and simultaneous frequency tagging of auditory and visual sources to study intermodal focused and divided attention in detection and discrimination performance. Divided-attention costs were smaller, but still significant, in detection than in discrimination. The auditory steady-state response (SSR) showed no effects of attention at frontocentral locations, but did so at occipital locations where it was evident only when attention was divided between audition and vision. Similarly, the visual SSR at occipital locations was substantially enhanced when attention was divided across modalities. Both effects were equally present in detection and discrimination. We suggest that both effects reflect a common cause: An attention-dependent influence of auditory information processing on early cortical stages of visual information processing, mediated by enhanced effective connectivity between the two modalities under conditions of divided attention.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Selective attention plays an important role in managing the continuous stream of incoming sensory information. By focusing attention on currently relevant information, the processing system is biased towards processing this information relative to currently less relevant and potentially conflicting or interfering information (Deco and Zihl, 2006; Desimone and Duncan, 1995). Attention can be focussed on a single source (e.g., object or location) but can also be divided across multiple information sources. These different modes of attention are referred to as focussed and divided attention.

Performance in divided-attention studies is usually interpreted in terms of a tradeoff between the relative amount of attention allocated to each task or source and the performance on each task. When a task is resource limited, performance is expected to improve as a monotonic function of the relative amount of attention allocated to the task. Performance tradeoffs between two concurrent tasks can be orderly represented in an "Attention Operating Characteristic" (AOC) which plots the joint performance on two tasks as a function of the attentional instructions (e.g. Bonnel and Hafter 1998, Sperling and Melcher, 1978; Norman and Bobrow, 1975).

Divided-attention costs can vary depending on both the features of the stimulus and the type of response to be made (Duncan, 1980). Bonnel et al. (1992) and Bonnel and Hafter (1998) showed that such costs can also depend on the nature of the task — concurrent target detection tasks did not yield statistically significant performance tradeoffs whereas strong tradeoffs were found for concurrent target discrimination tasks. Bonnel et al. (1992) and Bonnel and Hafter (1998) interpreted their results in terms of a hierarchical processing model (Hoffman, 1979; Duncan, 1981): Detection tasks depend on a relatively low level capacity-free processing stage, whereas discrimination additionally requires a higher level capacity-limited stage.

The formal distinction between detection and discrimination of targets has been debated in the literature (Brawn and Snowden, 2000; Kawahara et al., 2001). We follow the distinction made by Kawahara and collaborators, which seems most compatible with the procedures used by Bonnel et al. (1992) and Bonnel and Hafter (1998). They proposed that detection is involved when an observer is asked to distinguish between a uniform stimulus and a stimulus that contains a non-uniformity or discontinuity in space or time. Only the presence or absence of the non-uniformity is to be reported, not its identity. By contrast, discrimination (also referred to as identification) requires the observer to identify the nature of the discontinuity.

The notion that detection performance is not capacity-limited has not gone unchallenged. Several studies found target-detection performance to be enhanced when target location was cued in advance, suggesting that detection can be affected by attentional mechanisms (Muller and Humphreys, 1991; Hawkins et al., 1990; Downing, 1988; see also Lappin and Uttal, 1976). Bonnel et al. (1992) argued that such results do provide evidence that target detection can be affected by spatial uncertainty but do not undermine their proposal that detection performance is capacity free when, as in their own studies, spatial uncertainty has been removed.

* Corresponding author. Grote Kruisstraat 2/1, 9712 TS, Groningen, The Netherlands. Tel.: +31 50 363 6463.
   E-mail address: ritske.de.jong@rug.nl (R. de Jong).

In the present study, we followed the procedure of Bonnel and Hafter (1998), employing concurrent crossmodal auditory and visual detection or discrimination tasks with fixed stimulus locations and with several different attention allocation instructions. The tasks involved either detecting or discriminating a small and brief upward or downward change in tone volume in the auditory task and in stimulus brightness in the visual task. This particular procedure was chosen for several reasons. The data presented by Bonnel and Hafter (1998; Table 4) do seem to show effects of instructed attention allocation on performance even for concurrent detection. While these effects may have been due, at least in part, to participants matching performance to instructed relative attention (Navon, 1985), they might also reflect subtle but theoretically important capacity limitations in detection performance. Related to this point, participants may have become aware that different attentional policies impacted discrimination performance more than detection performance, and this may have led them to adopt less differentiating attention allocation policies in detection conditions. To address the possibility of different attentional strategies in crossmodal detection and discrimination, we recorded electroencephalographic (EEG) activity and used event-related potentials (ERPs) and, in particular, steady-state responses (SSRs) to gain information regarding the attentional mechanisms operative in the various experimental conditions.

The apparent absence of divided-attention costs in detection was reported by Bonnel et al. (1992) and Bonnel and Hafter (1998) for both intramodal and intermodal divided attention. In the present study, and as outlined more extensively below, we chose to focus on intermodal rather than intramodal attention in order to benefit from and contribute to the recent surge of research interest and findings regarding crossmodal effects in intermodal perception and attention (Driver and Noesselt, 2008; Macaluso and Driver, 2005; Stein and Stanford, 2008).

1.1. Using steady-state responses for studying auditory–visual intermodal attention

The auditory steady-state response (ASSR) refers to oscillatory EEG or MEG activity with constant amplitude and phase at the frequency of a periodic eliciting stimulus. The ASSR can be elicited by stimuli such as sequences of clicks, Gaussian tone pulses, or amplitude-modulated tones, and its amplitude is typically found to be maximal at stimulation frequencies around 40 Hz (Ross et al., 2005a). The ASSR is generally thought to result from periodic superposition of middle latency evoked responses (Galambos et al., 1981; Bohorquez and Ozdamar, 2008), but the alternative hypothesis that the ASSR reflects a separate neural oscillation that is driven by the periodic stimulus has also received empirical support (Ross et al., 2005a). Converging evidence from several neuroimaging studies and patient studies has shown the primary auditory cortex (medial portion of Heschl's gyrus) to be the main source of the 40-Hz ASSR (Gutschalk et al., 1999; Ross et al., 2002, Ross et al., 2005b; Simpson et al., 2005).

While the functional meaning of the ASSR remains unclear, an interesting hypothesis is that its amplitude may index sensory gain control in primary auditory cortex (Giard et al., 2000). Some studies have found effects of attention on the 40-Hz ASSR amplitude, for intramodal auditory selective attention (Bidet-Caulet et al., 2007; Skosnik et al., 2007; Gander et al., 2007) as well as for auditory–visual intermodal selective attention (Ross et al., 2004a,b; Saupe et al., 2009), but such effects have not always been found (Linden et al., 1987). A key feature of studies that obtained positive effects of attention on ASSR amplitude may be that they used tasks that required focused attention to the stimulus rhythm, or modulation frequency, eliciting the ASSR. Whether such effects can also be found when the stimulus rhythm is incidental to the task to be performed, and thus merely serves as a frequency tag for quantification of the

ASSR, is uncertain — one of the aims of the present study is to assess this possibility for the case of auditory–visual intermodal attention.

The visual counterpart to the ASSR is the visual steady-state response (VSSR; Regan, 1989), which is found with amplitude-modulated or flickering visual stimuli across a wide range of frequencies (Herrmann, 2001). In a recent elegant study, two main cortical sources of a 12-Hz VSSR were established: a main source in the primary visual cortex (V1) and a secondary source in the motion sensitive areas (MT/V5) (Russo et al., 2007). This suggests the possibility that VSSR amplitude may index sensory gain control in primary visual cortex. Positive effects of intramodal selective attention on VSSR amplitude have been reported in many studies that employed different paradigms such as visuospatial attention or selective color-based processing (e.g., Morgan et al., 1996; Muller et al. 1998a; Hillyard et al., 1998; Malinowski et al. 2007; Ding et al., 2006; Pei et al., 2002). In a visuospatial divided-attention task, VSSR amplitude for frequencies ranging from 8 to 23 Hz exhibited a graded dependency on attention, with substantially larger amplitudes for attended as compared to unattended locations and intermediate amplitudes when attention was divided across locations (Toffanin et al., 2009). In contrast to the many studies that used the VSSR to investigate neural mechanisms of intramodal visual attention, we are aware of only one study that used this measure to investigate effects of intermodal attention on visual information processing — Talsma et al. (2006) used the VSSR to demonstrate that attentional capacity between modalities is larger than that within one and the same modality.

1.2. The present study

ASSR amplitude and VSSR amplitude may provide useful indices of sensory gain control in early cortical auditory processing and visual processing, respectively. In the present study, we used these measures to investigate effects of intermodal auditory–visual selective or divided attention on early modality-specific cortical processing in detection or discrimination. The almost exclusive emphasis on steady-state responses may require some explanation, as we seem to have ignored the much more extensive and arguably better developed literature on the use of event-related potentials (ERPs) in the investigation of intramodal and intermodal attention (for review, see Driver and Noesselt, 2008; Macaluso and Driver, 2005). The reasons for this are mostly practical — in the present study we were unable to reliably measure short-latency and mid-latency ERP components (P50, C1, N1, P1, P2) that reflect sensory processing. Most likely, this absence of early ERP components was caused by the low saliency of the targets in both detection and discrimination.

The auditory stimulus consisted of a 40-Hz amplitude-modulated tone and the visual stimulus of a 24-Hz luminance-modulated colored square, both presented simultaneously for 2.5 s at approximately the same central location. Participants judged visual and auditory stimuli to share a common spatial source — thus, crossmodal selective attention could not be achieved by means of spatial selective attention to one or the other source of stimulation (Eimer and Schroger, 1998). Auditory and visual targets consisted of a brief (125 ms) decrease or increase of the mean loudness of the tone and of the main brightness of the colored square, respectively. Participants had to detect the presence of targets in the detection task, and had to identify the direction of change (decrease or increase) in the discrimination task. In different attention conditions, they were instructed to selectively attend to one and to ignore the other modality (100% auditory or 100% visual), or to divide their attention across modalities (20–80%, 50–50%, or 80–20% auditory–visual).

As relevant prior data regarding intermodal attention effects on ASSR and VSSR amplitudes in the literature are either inconsistent or largely absent (but see Talsma et al., 2006), the present study must be deemed exploratory. However, some possible outcomes will be briefly discussed. One possibility is that intermodal attention effects on ASSR and VSSR amplitude are mediated by largely the same neural

structures and mechanisms as those underlying intramodal attention effects (Talsma and Kok, 2001). In that case, graded effects of intermodal attention, similar to the graded effects of visuospatial attention on VSSR amplitude reported by Toffanin et al. (2009)), might be expected. Moreover, such effects might depend on whether the task involves detection or discrimination.

On the other hand, recent empirical and theoretical developments regarding crossmodal influences on modality-specific brain regions may suggest a very different outcome (for review, see Driver and Noesselt, 2008; Stein and Stanford, 2008), especially with regard to the distinction between divided-attention and focused-attention conditions. Many studies have found multisensory influences even in relatively low-order, early regions of sensory cortex (for review, Ghazanfar and Schroeder, 2006). Such influences may underlie multisensory effects on perceptual judgments, and are subject to constraints of spatial and temporal proximity or overlap between incoming signals from different sensory modalities (Macaluso and Driver, 2005). Possible mechanisms underlying crossmodal influences include rapid feedforward integration, thalamic influences, direct connections between primary cortices, and feedback from multisensory regions to sensory-specific brain areas (Driver and Noesselt, 2008). Apart from spatial and temporal constraints, which are likely to be fulfilled in the present study, crossmodal influences on modality-specific brain regions may also depend on attentional factors. Talsma et al. (2007) found crossmodal modulation of the auditory P50 only when participants attended to both audition and vision, and stated "We thus conclude that when attention is directed to both modalities simultaneously, auditory and visual stimuli are integrated very early in the sensory flow of processing (~50 ms poststimulus). Attention appears to play a crucial role in initiating such an early integration of auditory and visual stimuli. When only one modality is attended, the integration processes appear to be delayed." (p. 689). This idea suggests the interesting possibility of crossmodal effects on ASSR and VSSR amplitudes, where these effects may be particularly evident in the divided-attention conditions. More specifically, crossmodal interactions may manifest in two possibly related ways in divided-attention conditions. First, in these conditions 40-Hz activity, which specifically tags auditory input, may show up also at posterior occipital sites, whereas 24-Hz activity may show up at frontocentral, auditory sites. Second, crossmodal interactions may serve to enhance sensory gain in low-order auditory and visual cortices, thus suggesting the possibility that ASSR and VSSR amplitudes might actually be enhanced in divided-attention as compared to focused-attention conditions. Again, such effects may depend on whether the task involves detection or discrimination.

2. Experiment

2.1. Method

2.1.1. Participants

Ten healthy right-handed students (3 males) between the age of 19 and 23 (mean = 21.1; sd = 1.5) participated as paid volunteers. All volunteers had normal or corrected-to-normal vision. The study adhered to the Declaration of Helsinki and was approved by the local ethics committee. Informed consent was obtained from all participants. Due to technical malfunctioning, EEG data for one participant were lost.

2.1.2. Stimuli and apparatus

Visual stimuli were presented on a 17 inch CRT screen graphics (800 × 600 pixels with a 144-Hz refresh rate). Auditory stimuli were presented through a single loudspeaker placed at the top center of the computer monitor used for visual presentation; participants judged visual and auditory stimuli to share a common spatial source. The auditory stimulus was presented at 65 dB, measured at the location of the participants' ears. Stimuli were generated with the Matlab Psychtoolbox (Brainard, 1997; Pelli, 1997). Participants accommodated

their chin on a chin rest at a distance of 65 cm from the display monitor. The computer keyboard was used for the participants' responses.

The visual stimulus was a 4 × 4 cm (subtending 3.52º of visual angle) oscillating yellow square presented at the center of the screen. The luminance oscillation was created by changing the red and green values of the RGB values of the square according to a 24 Hz sine wave while the blue value of the RGB value was kept at zero. The mathematical description of the red and green values at time t is:

$$col(t) = b \cdot (1 + m \cdot \sin(2\pi \cdot f \cdot t)),$$

where b is the base value of the colors (125), m the modulation value (0.6), and f the modulation frequency (24 Hz). The visual target was an increase or decrease in the base value (b) for a period of 125 ms (three cycles of a 24 Hz sine wave). The auditory stimulus was a 500 Hz sine wave with 100% amplitude modulation at 40 Hz. The auditory target was an increase or decrease in the maximum volume of the tone for a period of 125 ms (exactly 5 cycles of a 40 Hz sine wave). To avoid any clearly perceptible discontinuities at target onset and offset, auditory target onset always coincided with the modulated amplitude being at or near its minimum value.

For both the visual and the auditory stimulus modulation onset was at $t = 0$. This was done to insure that the modulation was time/phase locked on every trial.

#### 2.1.3. Experimental design

All the participants performed both the detection and the discrimination tasks which were administered in separate sessions with a week in between sessions. Task order was counterbalanced across participants.

The event structure of a trial is illustrated in Fig. 1 (Panel A). A trial started with a presentation of a grey square in the center of the screen for 2000 ms — darkening of the square during the last 1000 ms indicated the upcoming presentation of the stimuli. After this warning period, the amplitude-modulated auditory and visual stimuli were presented for 2500 ms. During this 2500 ms period visual and auditory targets were presented. If a trial contained targets in both modalities, the targets were presented at the same time. In 80% of the trials, the targets were presented between 1000 and 2000 ms. In the other 20%, the targets were presented between 300 and 1000 ms. The latter trials served to force participants to remain focused during the entire interval and were excluded from analysis, ensuring a sufficiently long pre-target interval on the remaining analyzed trials for the SSR to develop and stabilize (Ross et al., 2005b).

In the detection task, participants were instructed to indicate whether a visual and/or an auditory target was present. Target probability for each modality was 50%, with a random half of targets being a decrease in volume or luminance and the other half an increase. Target presence and target type was randomly determined on each trial and independently for the two modalities.

In the discrimination task, auditory and visual targets were presented on each trial. Participants had to indicate if the target was a decrease or an increase in the volume of the tone or luminance of the square. Target type was randomly determined on each trial and independently for the two modalities.

Both tasks were administered in five different attentional conditions. Two conditions were focused-attention conditions where participants were instructed to focus their attention fully on one of the modalities (100% auditory or 100% visual). In the other three attention conditions they were instructed to divide attention between the two modalities in one of three different ways: 80% attention visual and 20% attention auditory, 50% visual and 50% auditory, and 20% visual and 80% auditory — in the 50–50% condition, they were told to divide attention equally between the modalities, whereas in the 80–20% they were told to focus most of their attention on one modality but not to completely ignore the other modality. Each of these conditions was administered in two consecutive blocks of 60 trials. The order of the attention
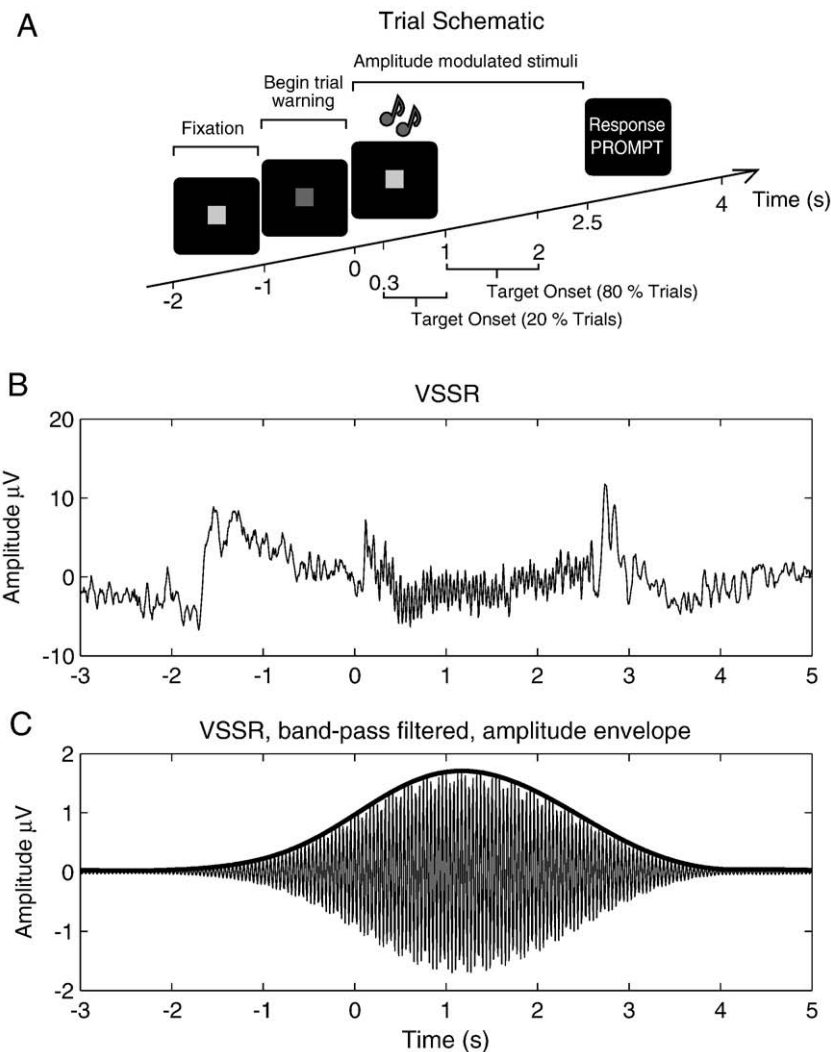
Fig. 1. Panel A. Schematic of trial procedure. A grey square was displayed for 1 s in the center of the screen and darkened for another second indicating the upcoming presentation of the 2.5 s amplitude-modulated auditory and visual stimuli. Targets could be presented in the visual and/or auditory modality. To ensure that participants deployed attention on the modulated stimuli for the whole trial duration, targets were presented between 1 and 2 s in 80% of the trials, or between 0.3 and 1 s in the remaining 20%. Only trials with targets between 1 and 2 s were analyzed. When targets were presented in both modalities they appeared simultaneously. Panels B and C depict the procedure for computing SSR amplitude. In this example VSSR amplitude is computed for a specific channel (PO3) in one of the conditions for an individual participant. Panel B: The visual steady-state response (VSSR) obtained by averaging across all available trials. Onset of the frequency tags is at $t = 0$, and an oscillatory EEG response can be seen to develop quickly and last until offset of the tags ($t = 2.5$). The large component at $t = -2$ is related to onset of the visual warning stimulus. Panel C: Visual steady-state response (VSSR) after narrow band-pass filtering ($f = 24$ Hz, $\sigma(f) = 0.2$ Hz). The amplitude envelope was obtained through Hilbert transform. Mean VSSR amplitude was computed across the interval 1–1.5 s after stimulus onset.

conditions was determined by a reduced order five Latin square in such a way that participant number six received the same condition order as participant number one. The order of the conditions was the same in the detection and the discrimination tasks for each participant.

Participants were prompted to respond after the 2.5 s period of stimulus presentation had expired. As Bonnel and Hafter (1998) found no effects of response order in a similar paradigm, participants were prompted to respond to the dominant modality first, by pushing one of two buttons labeled as 'no' and 'yes' in the detection task and as 'up' and 'down' in the discrimination task. In the two focused attention conditions the name of the attended modality served as the single prompt to respond. In the divided attention conditions participants were consecutively prompted to respond for each modality. The response for the dominant modality was made first in the 80/20 and the 20/80 conditions, and in the 50/50 condition response order was alternated between participants with, for each participant, opposite orders being used in the detection and the discrimination tasks. Responses for the auditory modality were made with the left hand and for the visual modality with the right hand.

2.1.4. Training

Each session started with a short explanation of the task to be performed, followed by a training that involved 100 trials of practice in each of the two focused attention conditions. Only stimuli and targets in the relevant modality were presented in these practice runs. During each practice run, a staircase tracking algorithm adjusted the target intensity (i.e. the volume difference between the baseline tone and the target tone or the luminance difference between the baseline square and the target square) to achieve an 80% accuracy level ($d' = 1.7$). The resulting target intensities were used in the experimental part of the session. Note that this procedure aimed to equate performance accuracies in the focused attention conditions across all combinations of task and modality.

2.1.5. EEG recording and electrode placement

EEG was recorded using a common reference amplifier: All channels were amplified against the average of all connected inputs. Electrical activity measured at the electrodes was amplified 20.000 times and digitally filtered using a low-pass FIR filter with a cutoff value of 140 Hz.

The EEG signals were recorded using Brain Vision Recorder (Brain Products GmbH, Munich, Germany). EEG signal was digitized at 500 Hz with a resolution of 0.02 μV.

A 70-electrode cap (Electro-cap International Inc, Eaton, Ohio, USA) was used for placement of the tin electrodes. All scalp positions in the International 10–20 System were used, with additional sites located midway between the 10–20 locations (Sharbrough et al., 1991) and six electrode positions 10% inferior to the standard parieto-occipital electrodes (FT9, PO9, O9, FT10, PO10, O10). Two additional electrodes were placed on the left and right mastoid for offline re-referencing of the EEG signal, and the ground electrode was placed on the sternum. Eye movements were monitored via bipolar recordings of the electrooculogram (EOG). For the horizontal EOG the electrodes were placed at the left and right external canthi and for the vertical EOG the electrodes were placed above and beneath the middle of the left eye. All inter-electrode impedances were kept below 5 kΩ for the EEG electrodes and 10 kΩ for the EOG electrodes.

2.1.6. Data analysis

Only trials where the target was presented after the first second of the stimulation interval were used for analysis (80% of the trials). This ensures that the SSR was well-developed and had stabilized by the time the target was presented.

2.1.7. Behavioral data

Dprime (d') values were calculated from hit and false alarm rates for each subject and each experimental condition, using standard procedures (Macmillan and Creelman, 2005):

$$d' = z(H) - z(F),$$

where $z(H)$ and $z(F)$ are the z-transformations of the hit rate and the false alarm rate respectively. Dprime values were subsequently subjected to repeated-measures analysis of variance (ANOVA). For all ANOVA's of behavioral and EEG data, Greenhouse–Geisser epsilon correction was used to adjust p-values, whenever applicable.

2.1.8. Electrophysiological data

All signals were re-referenced to the average of the mastoids. A 0.1 Hz high-pass filter (12 dB/oct slope) was applied to correct for slow wave fluctuations in signal amplitude. Ocular correction was performed using the method developed by Gratton et al. (1983)). All data was visually inspected for artifacts after segmentation.

2.1.9. ERPs

For the ERP analysis, segments of 1000 ms were made starting 200 ms prior to target onset on trials where responses were correct. A baseline correction was performed using the 200 ms interval preceding target onset. Segments were averaged and then smoothed with a 9 Hz low-pass filter (48 dB/oct slope). For the detection task, ERPs were calculated for each attention condition, separately for trials containing only a visual target, only an auditory target, or both targets. For the discrimination task, where every trial contained both a visual and an auditory target, ERPs were calculated for each attention condition. Because early components, including the N2, could not be reliably identified in the ERPs (see Fig. 3), presumably due to the use of the frequency tagging technique,[1] analysis was restricted to peak amplitude and latency values of the P300 component at Cz. These values were determined by simple peak picking in the 350–750 ms interval following target onset.

2.1.10. SSRs

The procedure used to estimate SSR amplitude is illustrated in Fig. 1 (Panels B and C). We used segments of 8000 ms that began

3000 ms before the start of the 2500 ms stimulation interval during which the amplitude-modulated auditory and visual signals were presented. Segments were averaged over trials, separately for each combination of task and attention condition, resulting in a steady-state potentials (SSRs; Fig. 1, panel B). Instantaneous SSR amplitudes (IAs) were then computed in two steps. First, SSRs were filtered, using FFT and a narrow Gaussian band-pass filter defined by a center frequency f (the driving frequencies of the visual and auditory stimuli, 24 or 40 Hz, respectively) and a standard deviation $\sigma(f)$ (0.2 Hz). Instantaneous amplitudes of the band-pass filtered SSRs were computed by means of (the absolute value of) the Hilbert transform (Fig. 1, panel C). This procedure is formally equivalent to computing instantaneous amplitude by means of a convolution with a complex Morlet wavelet[2] (for details, see Gladwin et al., 2006; Tallon-Baudry and Bertrand, 1999). SSR amplitude was computed as the average amplitude across the 1000–1500 ms interval following onset of the 2500 ms stimulation interval.

Average scalp distributions of ASSR and VSSR amplitudes were computed for each combination of task and attention condition. For each subject, normalized scalp distributions were computed by dividing the amplitude at each electrode by the maximum amplitude across all electrodes (McCarthy and Wood, 1985). Normalized amplitudes were then averaged across participants. Values for individual electrodes of the average scalp distributions vary between 0 and 1, with high values indicating consistently high relative SSR amplitudes across participants for that electrode and low values indicating consistently low relative SSR amplitudes.

The procedure used to estimate SSR amplitude as outlined above, i.e., time-domain averaging across trials followed by frequency-domain estimation of SSR amplitude, renders the nature of possible attentional effects on SSR amplitude ambiguous. Such effects could reflect veridical and consistent effects on SSR amplitude at the level of single-trials (i.e., veridical effects on sensory gain control) but could also reflect attentional effects on intertrial coherence of the phase relation between the driving oscillatory stimulation and the resulting EEG activity (Ding et al., 2006; Kim et al., 2007). To address this possibility, we computed SSR amplitude for single-trials, using the procedure detailed above, and then averaged these values across trials. By definition, average SSR amplitude estimated in this way cannot be sensitive to degree of intertrial phase coherence. Therefore, if attention primarily affects such phase coherence, attentional effects on SSR amplitude should no longer be present. In fact, attentional effects computed this way were virtually identical to those obtained with the former method of estimating SSR amplitude. We therefore conclude that the attentional effects on SSR amplitude, computed with the former method and to be reported below, almost certainly reflect genuine effects of attention on sensory gain control.
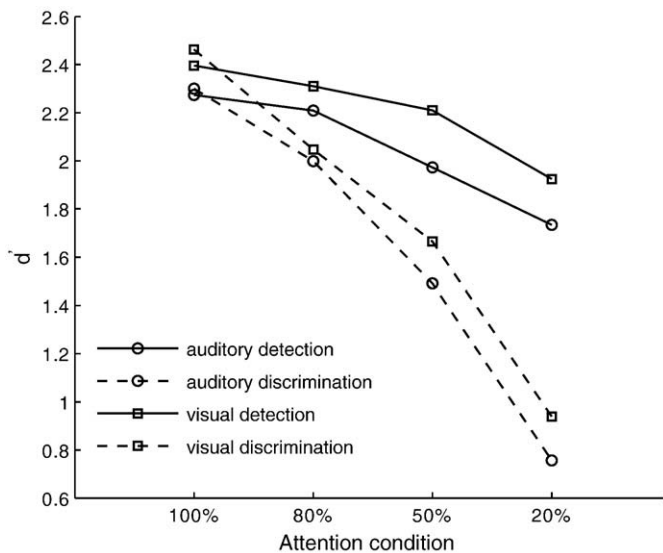
3. Results

3.1. Behavioral results

Averaged d'-values for each combination of modality (auditory, visual), task (detection, discrimination), and attention condition (100%, 80%, 50%, and 20% — note that no responses were given in the 0% condition) are presented in Fig. 2. Note that our procedure for

---

[1] Early ERP components also failed to become evident at any electrode position when the cutoff of the low-pass filter was set at 20 Hz.

[2] For the complex Morlet wavelet, frequency resolution ($\sigma(f)$) and time resolution ($\sigma(t)$) are reciprocally related as follows: $\sigma(f) \cdot \sigma(t) = 1/(2\pi)$. Small values for $\sigma(f)$ provide a desirably narrow focus on the driving or tagging frequency, but also result in a loss of temporal resolution in computing instantaneous amplitude. Because stimulation extended across a 2500 ms interval, during which SSR amplitude may be assumed to remain fairly constant (see Fig. 1), some loss of temporal resolution could be tolerated but up to a limit. Based on computer simulations, we determined that a value of 0.2 Hz for $\sigma(f)$ – corresponding to $\sigma(t) \approx 0.8$ s – provided a near-optimal compromise between the opposing requirements of frequency specificity and temporal specificity in the present study.

Fig. 2. Performance accuracies (d') as a function of attention condition for each combination of target modality (auditory, visual) and task (detection, discrimination). Attention condition refers to the proportion instructed attention to be allocated to the target modality (100% thus refers to the focused-attention conditions).

adjusting target intensity during training in order to achieve comparable levels of performance in the focused-attention (100%) condition for all combinations of modality and task, seems to have been quite effective. Divided-attention costs on performance were much stronger for discrimination than for detection, but some costs seem to be present also for detection. A $2 \times 2 \times 4$ (task $\times$ modality $\times$ attention) repeated-measures analysis of variance (ANOVA) yielded significant main effects of task, $F(1,8) = 6.4$, $p < .05$, and attention, $F(3,24) = 74.9$, $p < .001$, and a significant interaction of task and attention, $F(3,24) = 14.0$, $p < .001$. A separate analysis on the results for detection yielded a significant main effect of attention, $F(3,24) = 18.8$, $p < .001$.

3.2. Electrophysiological results

3.2.1. ERPs

ERPs at electrode Cz and time-locked to target onset are shown in Fig. 3, for all attention conditions in the detection and discrimination tasks. Note that these ERPs were computed from correct trials only. Separate panels show the ERPs associated with the three possible target events in the detection task: visual-only, auditory-only, and visual + auditory. The waveforms are dominated by the P300 that started after about 300 ms, whereas most of the earlier ERP components are hardly visible — as suggested earlier, this latter phenomenon is likely to result from the usage of frequency tagging in both modalities. Analysis of ERP waveforms was therefore restricted to peak P300 amplitude in the 400–800 ms interval.

The first analysis focused on P300 amplitudes for visual-only and auditory-only targets in the detection task. A repeated-measures ANOVA with target type (visual, auditory) and attention condition (recoded to indicate attention to the target type ranging from 0 to 100%) as within-subject factors yielded main effects of target type, $F(1,8) = 11.6$, $p < .01$, and attention condition, $F(4,32) = 18.3$, $p < .001$, and a significant interaction, $F(4,32) = 3.1$, $p < .05$. Pairwise comparisons between the attention conditions showed no significant differences between the 100%, 80%, and 50% conditions, significant differences between each of these three conditions and both the 0% and 20% conditions, and a significant difference between the 0% and 20% conditions. Thus, visual targets elicited larger P300 components than auditory targets, and, with the exception of the 20% condition, the effect of attention on P300 amplitude was approximately all-or-none.

Inspection of Fig. 3 suggests that P300 amplitudes for bimodal targets in the detection task may roughly correspond to the sum of amplitudes for visual-only and auditory-only targets in that task. An analysis designed to test for possible differences between amplitudes for bimodal targets and the sum of unimodal targets yielded no significant differences. This result is consistent with the notion that the P300 response to bimodal targets in detection is composed of the linear sum of independent P300 responses to each of the constituent unimodal target events (Teder-Salejarvi et al., 2002; Talsma and Woldorff, 2005).

P300 amplitudes for bimodal targets in detection and discrimination were analyzed in a repeated-measures ANOVA with task and attention condition as within-subject factors. Apart from the main effects of attention condition, this analysis yielded as a new result a borderline significant main effect of task, $F(1,8) = 4.5$, $p < .07$, reflecting somewhat smaller P300 amplitudes when the task required discrimination as compared to detection.

3.2.2. VSSR

Averaged topographical maps of the 24 Hz VSSR amplitude are shown in Fig. 4, for all combinations of task and attention condition. These maps show a pronounced and consistent maximum at the most posterior electrodes. Also evident, for detection and discrimination, are the increased values at these electrodes for divided-attention as compared to the focused-attention conditions. Statistical analysis of VSSR amplitude was based on the average amplitudes computed across the four electrodes where this amplitude, averaged across all task by attention condition combinations, was maximal. This subset of electrodes was determined separately for each participant, but showed a high degree of consistency across participants, containing only electrodes at posterior occipitoparietal locations. Estimates of the signal-to-noise ratio (SNR) of VSSR amplitudes that were used in statistical analysis were computed in the following way. We used exactly the same procedure as that used for computing the 24 Hz response, but now centered the Gaussian filter at 22 Hz or 26 Hz. Dividing the 24 Hz amplitude estimate, averaged across all task x attention condition combinations, by the average of the 22 and 26 Hz estimates, similarly averaged across experimental conditions, provides an estimate of the frequency specificity, or SNR, of the 24 Hz VSSR amplitude. Values of this estimate across participants ranged from 11.5 to 48.3, with an average of 28.4, indicating excellent frequency specificity for VSSR amplitude estimation.

To accommodate the substantial individual differences in VSSR amplitude, values were normalized for each participant by dividing by the average value across all ten (task x attention condition) conditions. Fig. 5 shows normalized mean VSSR amplitude as a function of task and attention condition. For both detection and discrimination, VSSR amplitude increased markedly from focused-attention to divided-attention conditions, culminating at the 50–50% condition. Amplitudes also appear to be slightly larger in the 100%-visual as compared to the 100%-auditory condition. A repeated-measures ANOVA with task and attention condition as within-subject factors yielded only a main effect of attention condition, $F(4,32) = 8.5$, $p < .001$. Polynomial contrast analysis showed this effect to be explained completely by the quadratic contrast, $F(1,8) = 24.2$, $p < .001$ — the linear contrast did not approach significance, $F(1,8) = 1.5$, $p > .26$.

To assess the relation between VSSR amplitude and performance accuracy, VSSR amplitude was computed separately for correct trials and incorrect trials, and separately so for visual targets and auditory targets and for detection and discrimination. In order to have a sufficient number of trials for stable estimates of VSSR amplitude, errors were pooled across attention conditions — corresponding 'pooled' values for correct trials were obtained by weighted averaging of the amplitudes for individual attention conditions, with the weights corresponding to the relative frequency of errors in each condition. For each participant, the resulting values were normalized by dividing by the average value of
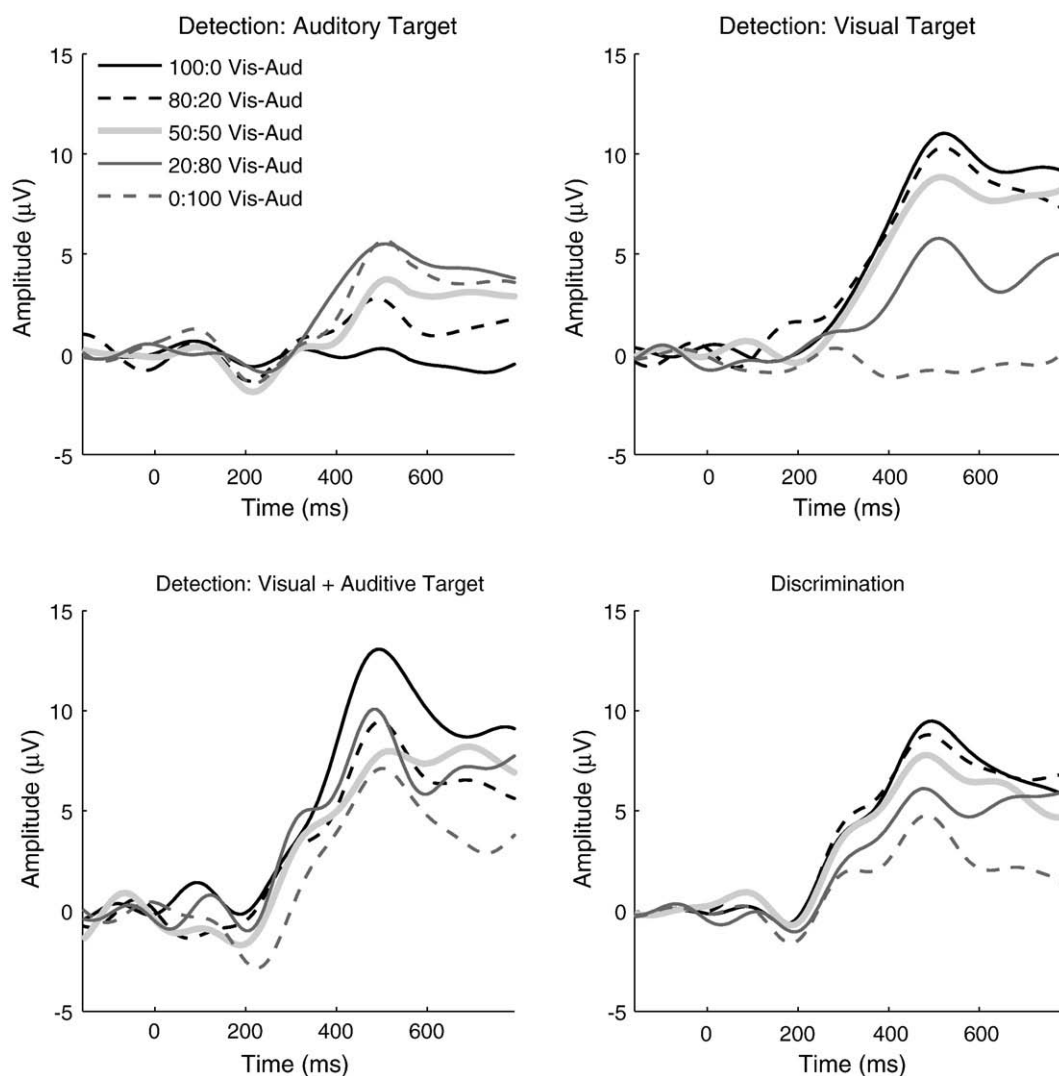
Fig. 3. ERP waveforms at Cz. Target onset is at $t = 0$. The upper and lower left panels display the waveforms in each attention condition separately for each of the three possible target types in the detection task (auditory-only, visual-only, auditory + visual). The lower right panel shows the waveforms for the discrimination task.
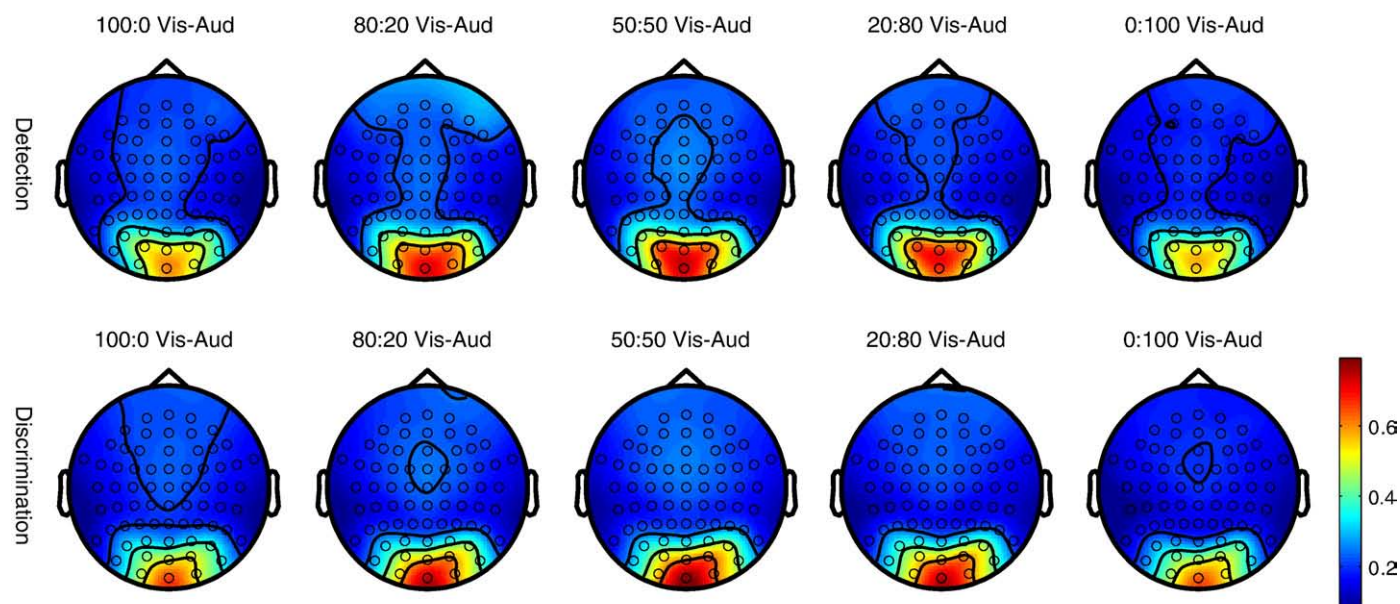
Fig. 4. Topographic maps of the visual steady-state response (24 Hz). The upper panel depicts separate maps for each attention condition in the detection task; the lower panel does so for the discrimination task. Higher values represent higher average normalized amplitudes across participants at that location.
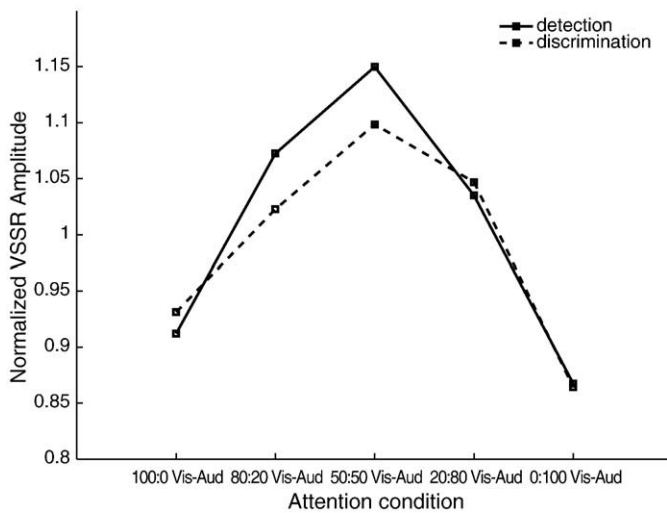
Fig. 5. Average normalized amplitudes of the visual steady-state response, for all attention conditions and separately for detection and discrimination.

'correct' amplitudes across the four combinations of task and target modality. The results of this analysis, averaged across participants, are shown in the left panel of Fig. 6. Amplitudes seem somewhat smaller for detection as compared to discrimination. Amplitudes for auditory targets appear to be somewhat larger for errors than for correct responses in both detection and discrimination, whereas an opposite but more pronounced effect seems to be present for visual targets. A repeated-measures ANOVA with correctness, task, and target modality as within-subject factors yielded a significant main effect of correctness, $F(1,8) = 9.0$, $p < .02$, and a significant interaction of correctness and target modality, $F(1,8) = 10.9$, $p < .015$. Separate analyses for visual and auditory targets yielded a significant main effect of correctness for visual targets, $F(1,8) = 22.2$, $p < .002$, but not for auditory targets, $F(1,8) = 4.1$, $p > .10$.

3.2.3. ASSR

Averaged topographical maps of the 40 Hz ASSR amplitude are shown in Fig. 7, for all combinations of task and attention condition. Replicating previous results (e.g., Saupe et al., 2009), these maps show ASSR amplitude to be largest across a fairly wide frontocentral region. As for the VSSR, statistical analysis of ASSR amplitude was based on the average value computed across the four electrodes where this amplitude was maximal — all 36 electrodes (9 participants * 4 electrodes) used in this computation had a frontocentral location. Following the same procedure as described above for VSSR (24 Hz) amplitude, but now with 38 and 42 Hz as the bracketing frequencies, yielded for the 40-Hz ASSR amplitude an SNR range of 2.6 to 10.0, with an average of 5.5. These values indicate acceptable and useful frequency specificity for ASSR amplitude.

ASSR amplitudes were normalized, using the same procedure as that described earlier for VSSR amplitude. A repeated-measures ANOVA with task and attention condition as within-subject factors yielded no significant effects (all p's > 0.17).

A possible relation between ASSR amplitude and accuracy for auditory and visual targets was analyzed in the same way as previously outlined for VSSR amplitude. The results of this analysis, averaged across participants, are shown in the right panel of Fig. 6. Amplitudes were somewhat smaller for detection as compared to discrimination. More importantly, amplitudes for auditory targets appeared to be smaller for errors than for correct responses in both detection and discrimination, whereas the opposite effect seems to be present for visual targets. A repeated-measures ANOVA with within-subject factors correctness, task, and target modality yielded as the only significant effect an interaction of correctness and target modality, $F(1,8) = 10.1$, $p < .02$. Separate analyses at each level of target modality failed to show significant effects of correctness (p's > .13). Thus, these results suggest an opposite but weak relation between ASSR amplitude and performance accuracy for auditory targets and visual targets. Note that these latter effects are almost the mirror image of those found for VSSR amplitude depicted in the left panel of Fig. 6.
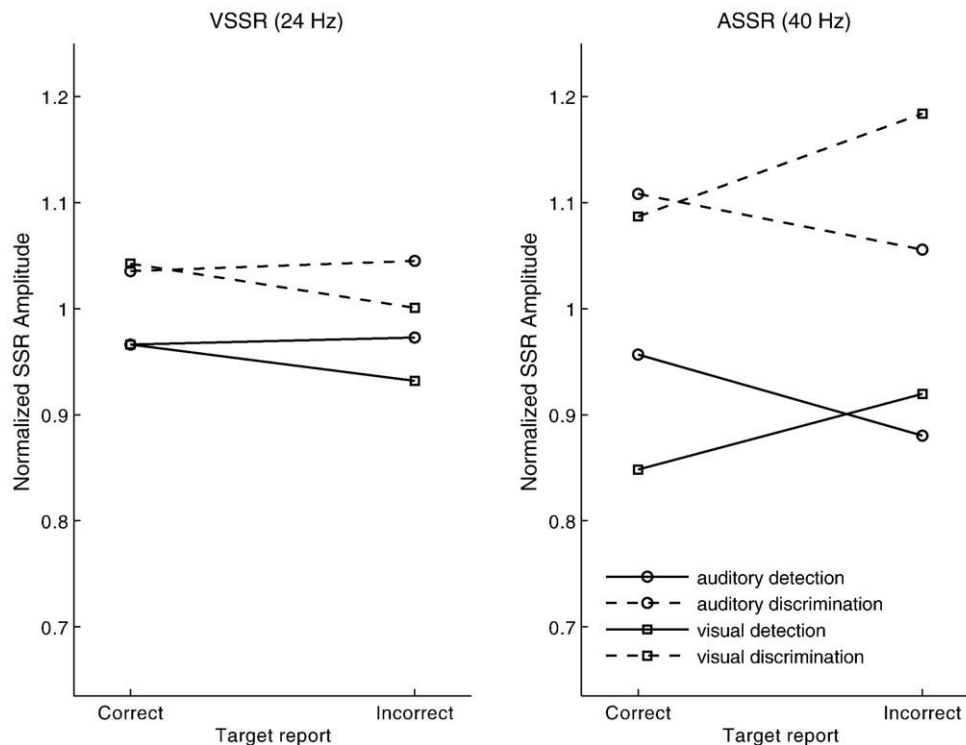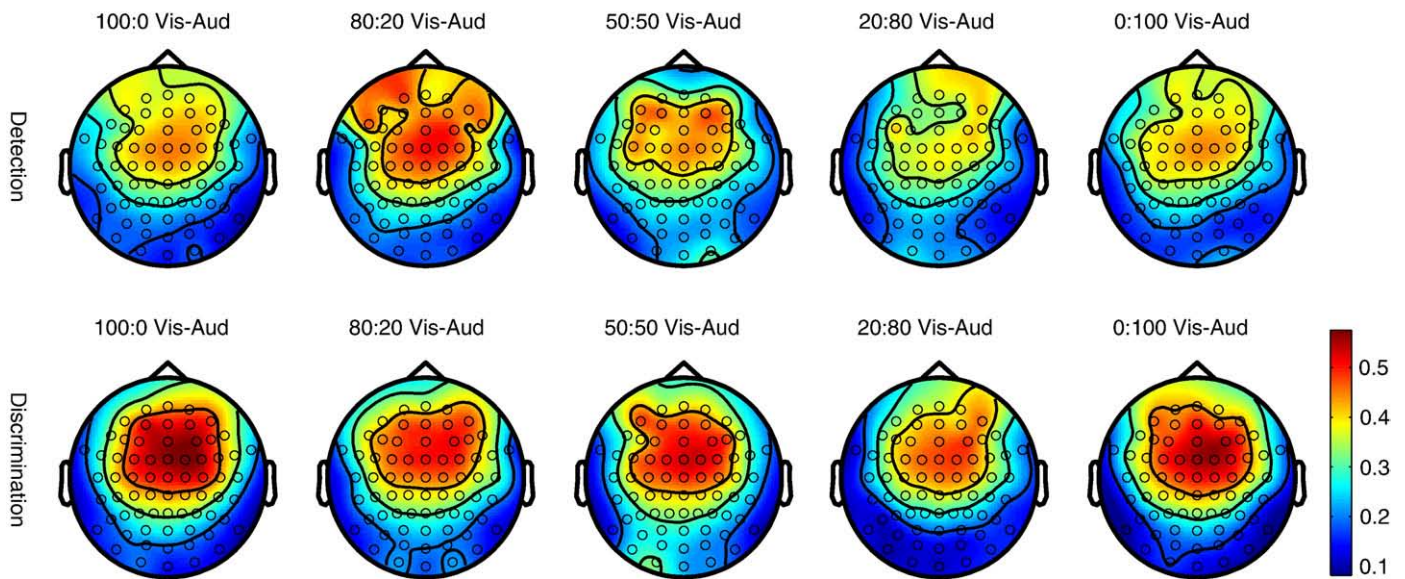

Fig. 6. Normalized amplitudes of the visual steady-state potential (VSSR; left panel) and auditory steady-state potential (ASSR; right panel), as a function of correctness of the target response, target type (auditory, visual), and task (detection, discrimination).

Fig. 7. Topographic maps of the auditory steady-state response (40 Hz). The upper panel depicts separate maps for each attention condition in the detection task; the lower panel does so for the discrimination task. Higher values represent higher average normalized amplitudes across participants at that location. Note the emergence of the 40 Hz response at posterior locations during the 50–50% divided-attention condition.

Close inspection of the topographical maps in Fig. 7 reveals a slightly enhanced presence of 40 Hz activity at posterior occipital electrodes for the 50–50% divided-attention condition in both detection and discrimination. Because such an effect would be of considerable theoretical significance but the observation is post hoc, we analyzed it in detail. We averaged ASSR amplitude across the three most posterior electrodes (O9, Iz, O10), as the effect seems most prominent at those electrodes — very similar results were obtained when we used the same set of four electrodes used for computing VSSR amplitude. Because the amplitudes involved are very small (less than 0.1 μV) and to accommodate the very substantial individual differences, we also averaged amplitudes across detection and discrimination to enhance robustness and then log-transformed them (using the natural logarithm). Because the posterior electrodes are prone to EMG artifacts, and the 50–50% divided-attention condition might be considered the most challenging one and lead participants to tense up, it is possible that the effect represents an EMG artifact. In addition, the effect might reflect enhanced gamma activity in background EEG at posterior sites under more challenging conditions (Fell et al., 2003). To explicitly check for these alternative explanations, we used exactly the same procedure used to compute 40 Hz amplitudes to compute amplitudes at two neighboring frequencies, 38 and 42 Hz, and then averaged across these two frequencies. As both EMG activity and gamma-band EEG activity have broad spectra, we reasoned that such activities should be equally evident at those neighboring frequencies, thus providing an appropriate check for the possibility that such activities might underlie the 40 Hz effect at posterior sites.

Average (log-transformed) ASSR (40 Hz) amplitudes and corresponding mean 38–42 Hz amplitudes are presented in Fig. 8, as a function of attention condition. The ASSR amplitude is slightly larger than the 38/42 Hz control value at the focused-attention conditions, but this difference is substantially increased at the divided-attention conditions and is maximal at the 50–50% condition. The relatively small attention effects on the 38/42 Hz control amplitude that can be seen might reflect a modest contribution of EMG or gamma-band activity to the attentional effects on ASSR amplitude, but might also reflect leakage of 40 Hz activity to neighboring frequencies due to nonlinear neural interactions. Repeated-measures ANOVA with component (40 Hz, 38/42 Hz) and attention condition as within-subject factors yielded main effects of component, $F_{(1,8)} = 14.0$, $p < .01$, and attention condition, $F_{(4,32)} = 7.3$, $p < .005$, and a significant interaction

of these factors, $F_{(4,32)} = 5.4$, $p < .01$. Polynomial contrast analysis revealed that the latter effects involving attention condition were almost completely explained by the quadratic contrast ($F_{(1,8)} = 13.8$, $p < .01$, for the main effect of attention condition, and $F_{(1,8)} = 11.6$, $p < .01$, for the interaction). Analysis of simple effects showed that the effect of attention was significant for the 40 Hz component ($p < .001$), but not for the 38/42 Hz component ($p > .15$). These results seem to provide compelling evidence that the enhanced presence of 40 Hz activity at posterior sites in divided-attention conditions represents genuine neural activity at occipital sites that is specifically associated with the auditory input stream.

Inspired by this latter effect, we also undertook several analyses aimed towards identifying possible attention-dependent VSSR (24 Hz) activity at frontocentral electrodes, which could not be explained in terms of volume conduction from the primary occipital sources of the VSSR. None of these produced even a hint of a possible presence of such effects.
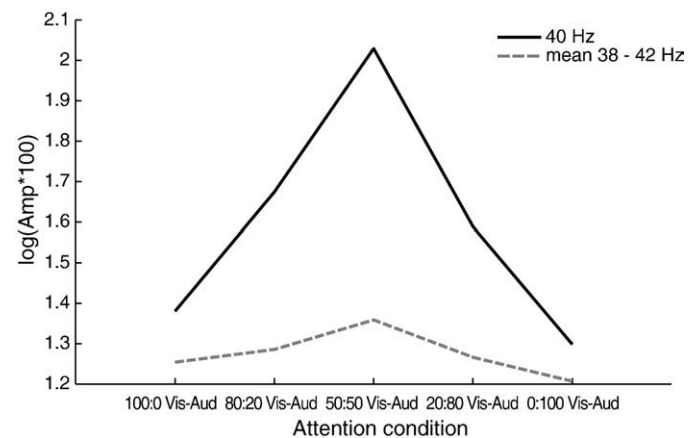
Fig. 8. Mean log-transformed amplitude of the 40 Hz response at posterior occipital locations, as a function of attention condition. The amplitude of the 38/42 Hz response serves as an important control to assess possible influences of EMG artifacts and gamma-band EEG activity. See text for further explanation.

4. Discussion

In the present study we explored the effects of bimodal divided attention on performance during auditory–visual detection and discrimination, using a paradigm similar to that developed by Bonnel and Hafter (1998). The behavioral results showed clear dual-task costs when attention had to be divided between the auditory and visual modalities for discrimination performance and also, though the costs were substantially smaller, for detection performance. These results replicate earlier results by Bonnel et al. (1998) who, as we argued before, also seem to have obtained divided-attention costs in detection performance when, as in the present study, target intensity was adjusted in such a way that overall performance levels were approximately equated for the focused-attention conditions in the detection and discrimination tasks. Thus, the present evidence fails to support the notion that auditory–visual bimodal detection and discrimination are mediated by qualitatively different processing stages, a pre-attentive capacity free stage in detection and an attention-dependent capacity limited stage in discrimination (Bonnel et al., 1992; Bonnel and Hafter, 1998). Instead, the evidence seems more consistent with the notion that bimodal detection is in fact subject to capacity limitations but that these limitations are generally less severe than those delimiting dual-task performance with auditory–visual discrimination tasks.

Results from a study by Kawahara et al. (2001), who used an attentional-blink paradigm, further substantiate this view. They found an attentional blink (i.e., impaired detection of the second target when it followed the first target within 100 to 500 ms — see Raymond et al., 1992, for an introduction to the paradigm) when the second target had to be identified but not when it had to be detected — an attentional blink in detection performance was obtained only when the stimuli were highly degraded. These results also suggest that detection is subject to capacity limitations but that such restrictions become manifest in performance only when detection is sufficiently difficult in terms of signal-to-noise ratios.

The fact that divided-attention costs were much smaller for detection than for discrimination, even when performance in the two tasks was equated in the focused-attention conditions, might be explained in terms of a steeper performance-resource tradeoff for discrimination than for detection (Norman and Bobrow, 1975). In the present study we sought to address an alternative, but compatible, account — that people may tend to adopt less differentiating attention allocation policies in detection conditions, perhaps because they become aware that different allocation policies may impact performance in concurrent detection less than that in concurrent discrimination. We will discuss the results for the various electrophysiological measures in detail below, but the fact that attentional effects on all these measures were remarkably similar for detection and discrimination, fails to provide any support for this alternative account.

4.1. Event-related potentials

Because the various early-onset ERP components were either absent in the present study or too small to allow for useful analysis, presumably due to the low salience of the targets, only the P300 component could be considered. In the detection task, mean P300 amplitude, which was computed across correctly detected targets only, was larger for visual-only targets than for auditory-only targets. More important, P300 amplitude on single-target trials was found to scale systematically with instructed attention allocation — as more attention was allocated to one modality, P300 amplitude increased for targets in that modality while it decreased for targets in the other modality. These results are in agreement with findings by previous studies that also showed an increase of P300 amplitude as a function of attention allocation (e.g. Mangun and Hillyard, 1990; Hoffman et al., 1985, Wickens et al., 1983; Isreal et al., 1980). Such results have

been interpreted in terms of P300 amplitude reflecting the resource distribution of a limited capacity processing system (Hoffman et al., 1985). However, others have argued that P300 amplitude is closely tied to the behavioral outcome, and might reflect the effects of attentional selection rather than the attentional modulation and selection of relevant information itself, with the latter causal processes being reflected by earlier ERP components such as the P1 and N1 (Mangun and Hillyard, 1990; Toffanin et al., 2009). Consistent with this latter notion, the effects of relative attention allocation on P300 amplitude were distinctly nonlinear, with relatively large and similar amplitudes in the 100%, 80%, and 50% conditions, a smaller amplitude in the 20% condition, and an essentially absent P300 in the 0% condition. Thus, it is questionable whether the attentional effects on P300 amplitude can provide any more direct information on the allocation of attention across the two modalities than what is provided by effects on overt behavior. Interestingly, P300 amplitude for bimodal targets closely equaled the summed amplitudes for unimodal targets — this might suggest that at the level of processing indexed by the P300, there is little or no overhead cost associated with the requirement to divide attention in bimodal detection (Bonnel et al., 1992; Bonnel and Hafter, 1998).

In the discrimination task, targets in the different modalities were always presented simultaneously. Consistent with the larger amplitudes for visual-only targets than for auditory-only targets in the detection condition, P300 amplitude in the discrimination condition steadily increased as more attention was allocated to the visual modality. When compared to the P300 amplitudes for bimodal targets in the detection condition, amplitudes in the discrimination condition were somewhat smaller. As bimodal targets were far less common in the detection condition, this latter effect may reflect the well-known effect of probability on P300 amplitude (Johnson, 1984, 1986).

4.2. Frequency tagging: ASSR and VSSR

Several aspects of the results from frequency tagging were unexpected and, we suggest, quite remarkable. First, ASSR amplitude (40 Hz) at frontocentral electrodes, where the amplitude exhibited a clear maximum, showed no significant effects of intermodal attention. The lack of a difference between the focused-auditory and focused-visual conditions replicates earlier findings (Linden et al., 1987) and reinforces the notion that such a difference may be found only when the target event is defined in terms of a change in the modulation frequency itself (Ross et al., 2004a, 2004b; Saupe et al., 2009; Skosnik et al., 2007). ASSR amplitude fluctuates from trial to trial. Such fluctuations were found to predict performance accuracy, which high ASSR amplitude being associated with accurate performance with auditory targets but with inaccurate performance with visual targets. Though these results are only correlational, it is tempting to infer a causal relationship where ASSR amplitude indexes relative preparedness to process auditory as compared to visual information. The contrast with our failure to find effects of intermodal attention instructions on ASSR amplitude may be more apparent than real. The correct–incorrect distinction may simply provide a more powerful contrast than that provided by the various attention conditions – after all, different attention instructions had significant but relatively mild effects on performance accuracy. Furthermore, even the correct–incorrect contrast failed to produce significant effects on accuracy when visual targets or auditory targets were considered separately. The combined results may point to a subtle, but theoretically significant, relation between (EEG-based) ASSR amplitude and relative processing bias towards auditory versus visual sources of information, which requires considerable power to detect. This perspective makes it understandable why and how such power is enhanced when the target is defined in terms of a change in the tagging frequency itself Ross et al. (2004a,b). Our only reservation with this solution to boost power is that it would seem to do away

with the very elegance and promise of frequency tagging as a technique that can provide a continuous window on perceptual information processing in the brain without influencing or interfering with such processing.

Clear attention effects on ASSR amplitude were found at posterior occipital sites, where the 40 Hz response was markedly stronger in divided-attention as compared to focused-attention conditions. It is important to note that these effects were very small in absolute size, approaching, but not exceeding, the resolution limits of our EEG recording system. These effects are unlikely to reflect either EMG artifacts or gamma-band EEG activity, as they were not found, or were much smaller, at neighboring frequencies (38/42 Hz). Neither can they be due to passive volume conduction from primary ASSR sources in auditory cortex, as no significant attention effects were found there. By exclusion then, as the 40 Hz frequency specifically tags the auditory input stream, the enhanced 40 Hz response at occipital sites in divided-attention conditions must be taken to reflect an enhanced influence of auditory input or auditory processing on neural activity in occipital cortex. To our knowledge, this is the first demonstration of an unambiguous attention-dependent influence of auditory information on occipital activity using EEG (Talsma et al., 2007, may have demonstrated such an influence of visual information on auditory regions). The nature of such an enhanced influence in divided-attention conditions remains to be clarified. The fact that ASSR amplitude at anterior-central sites showed no attention effects would seem to rule out the possibility that the effects at posterior sites are due to a stronger neural drive (as opposed to passive volume conduction) from auditory regions, presumably primary auditory cortex, in divided-attention conditions. Rather, these effects might reflect stronger effective connectivity from auditory to visual cortical areas in divided-attention conditions. That issue will be discussed in more detail later.

VSSR amplitude (24 Hz) exhibited a clear maximum at posterior occipital sites, with a highly favorable signal-to-noise ratio. Attention effects on posterior VSSR amplitude resembled those found for ASSR amplitude at posterior sites, with higher amplitudes obtained in divided-attention conditions as compared to focused-attention conditions and with both effects being fully accounted for by the quadratic contrast in polynomial contrast analyses. We stress that almost identical attention effects were obtained when VSSR amplitude was computed in a way that rendered it completely insensitive to possible attention-related differences in intertrial phase coherence (see Method). Thus, these effects on VSSR amplitude almost certainly reflect genuine effects on sensory gain control (Kim et al., 2007).

VSSR amplitude in the focused-visual condition did not differ from that in the focused-auditory condition. This lack of an effect of intermodal selective attention contrasts strongly with the many positive findings of clear intramodal visuospatial attention effects on VSSR amplitude, also in studies that used tagging frequencies similar to the 24 Hz used in the present study (Muller and Hillyard, 2000; Muller et al. 1998a,b; Morgan et al., 1996; Toffanin et al., 2009). The latter findings have generally been interpreted in terms of modulation of sensory gain control in primary or secondary visual cortex, and this interpretation is corroborated by evidence from numerous fMRI studies of visuospatial attention (Kastner et al., 1998; Russo et al., 2003; Tootell et al., 1998). Several recent fMRI studies of auditory–visual intermodal attention have also found evidence for enhanced activity in modality-specific cortical areas associated with the attended modality in conjunction with diminished activity in cortical areas associated with the unattended modality (for a recent example and overview, see Johnson and Zatorre, 2005, 2006). Importantly, these cortical areas are typically found not to include primary auditory and visual cortex (e.g., Johnson and Zatorre, 2005, 2006, Shomstein and Yantis, 2004). As occipital VSSR has been associated with sources in primary visual cortex (Russo et al., 2007), the fact that evidence from fMRI studies indicates that effects of selective intermodal attention are largely restricted to secondary or higher visual cortical areas, may thus be consistent with the lack of selective intermodal attention on VSSR amplitude in the present study.[3]

As noted before, the enhanced VSSR amplitude in divided-attention conditions resembles the results for ASSR amplitude (40 Hz) at occipital sites, and we suggest that both effects may share a common cause: Enhanced effective connectivity from auditory to visual cortex in divided-attention conditions so that activity in auditory cortex can affect, directly or indirectly, activity in visual cortex during crossmodal auditory–visual attention. The central idea here is that enhanced 40 Hz activity at occipital sites serves the critical function of unambiguously indexing such enhanced effective connectivity, but that 40 Hz activity itself constitutes only a small portion of total neural activity in auditory cortex that, once effective connectivity has been established, can exert a more general influence on neural activity levels and patterns in visual cortex. If VSSR amplitude is assumed to reflect sensory gain control in primary visual cortex (Russo et al., 2007), the present findings would seem to provide strong evidence for an influence of auditory information processing on early cortical stages of visual information processing. Possibly related small but reliable changes in occipital activity due to auditory-spatial attention were recently reported by Wu et al. (2007). More speculatively, such an influence may have been responsible for results by van der Burg et al. (2008) who demonstrated that a nonspecific auditory beep could critically facilitate the detection of low-salience visual targets. Finally, such an influence on sensory gain in visual cortex may have been responsible for the finding of Talsma et al. (2006) that attentional capacity between modalities (auditory and visual) is larger than that within the visual modality alone.

An alternative interpretation of the enhanced VSSR amplitude in divided-attention conditions, is that VSSR amplitude depends on general arousal level and that this level was higher in the presumably more challenging divided-attention conditions. One problem with this interpretation is that it fails to explain why, as discussed earlier, VSSR amplitude is not enhanced in divided-attention conditions in intramodal studies. Also, without a specification of the neural systems and mechanisms underlying possible effects of general arousal on VSSR amplitude, this interpretation may be difficult to distinguish from interpretations that account for crossmodal effects in terms of neural interactions at midbrain structures (e.g. thalamus) that influence cortical processing (Driver and Noesselt, 2008). Such a distinction could perhaps be made in future studies by manipulating the relative positions of the auditory and visual sources, as crossmodal interactions are typically dependent on close spatial vicinity of sources (Driver and Noesselt, 2008), whereas the difficulty of divided-attention conditions, and thus general arousal level, should instead be greater when sources are spatially separate. For now, we prefer the account in terms of enhanced effective connectivity over that in terms of general arousal as the former but not the latter can also readily account for the presence of the ASSR at occipital sites in divided attention conditions.

In the present study, we failed to obtain any evidence for a reverse attention-dependent influence from visual information processing on auditory information processing. This might point to a truly asymmetric interaction between vision and audition under the conditions of the present experiment, or it could be due to limitations of the measures and techniques we used to find evidence for such a reverse influence. Seemingly consistent with the former possibility, Giard and Peronnet (1999), in what could arguably be considered an auditory–visual divided-attention task, found influences of auditory information on occipital activity to be both earlier (~50 ms) and stronger than the reverse influences.

---

[3] Some previous studies using the frequency tagging technique have found the second harmonic response to be particularly sensitive to attentional manipulations (e.g., Pastor et al., 2007). For this reason we also checked for possible effects of selective intermodal attention at 48 Hz — no such effects were found.

It is important to point out that the effects of intermodal attention on VSSR amplitude were very similar for detection and discrimination. Apart from serving as a useful indication of the robustness of the present results, this similarity also serves to indicate that attentional strategies and mechanisms are likely to have been largely the same in detection and discrimination. Attention-dependent crossmodal links of the type demonstrated here might have important functions in supporting flexible multisensory integration. The tasks we used in the present study, with target events in one modality being independent of and unrelated to those in the other modality, provided no basis for useful multisensory integration, begging the question of whether and how such links may have influenced task performance. We have interpreted VSSR amplitude to reflect sensory gain at low levels of cortical visual processing, but the positive functional connotations of the concept of sensory gain may be reason to reconsider that interpretation. As noted earlier, recent evidence tends to support the hypothesis that the ASSR reflects a separate neural oscillation that is driven by the periodic stimulus (Ross et al., 2005a). If we apply the same hypothesis to the VSSR, it becomes possible to see how input from auditory processing regions, even when it carries no useful information – and thus should be considered noise – may raise the level of background activity in early visual cortex and thereby enhance VSSR amplitude as background activity is entrained by, or feeds, the neural oscillation. However, a higher level of background activity, or noise, would not necessarily facilitate the processing of discrete visual events, and would perhaps be more likely to actually hamper such processing. Thus, attention-dependent crossmodal links, by enabling crosstalk, may have contributed to divided-attention costs in performance in the present study (Navon and Miller, 2002). The more important and general point is that tasks, such as the ones used here, that provide little or no possibilities for useful multisensory integration, may nevertheless be powerful tools to establish potential neural mechanisms underlying such integration. To assess the functional significance of such mechanisms requires the complementary use of tasks or paradigms that allow for useful types and degrees of multisensory integration.

4.3. Multisensory integration: ERPs and SSRs

The SSR is easily quantified in the frequency domain and knowledge about the neural generators of the ASSR and VSSR greatly enhance their utility in assessing neural activity in specific cortical regions. Two major sources of the VSSR have been identified in primary visual cortex and in the motion-sensitive area MT/V5 — in addition, activity in mid-occipital and ventral occipital region appear to make minor contributions to the VSSR (Russo et al., 2007). Together with the posterior occipital distribution of the VSSR in the present study, this has led us to interpret VSSR amplitude to index neural activity in early regions of visual cortex. While several studies have identified primary auditory cortex as the main source of the ASSR, a recent study analyzed sources by means of PET and found six sources of the ASSR, including temporal, frontal, and parietal regions as well as regions in the brain stem and cerebellum (Reyes et al., 2005). Given the evidence, it seems reasonable to associate ASSR amplitude primarily with neural activity in early regions of auditory cortex.

The regional specificity of ASSR and VSSR in conjunction with the fact that their temporal resolution is inherently poor, imply that these measures do not allow for a chronometric analysis of brain activity across different cortical areas. Several possible mechanisms underlying crossmodal influences have been proposed, including rapid feedforward integration, thalamic influences, direct connections between primary sensory cortices, and feedback from multisensory regions to sensory-specific brain areas (Driver and Noesselt, 2008). We believe that the present results, while providing clear evidence for attention-dependent crossmodal influences, are mute regarding the issue of underlying mechanisms, because of this inherent limitation of the SSR measure. In contrast, the high temporal resolution of ERPs combined with the high spatial resolution of fMRI, can provide a picture of the sequential activation of cortical areas during perceptual processing. Some interesting results that demonstrate the potential of this technique and speak to underlying neural mechanisms are already available, for both intramodal attention (e.g., Russo et al., 2003) and intermodal attention (e.g., McDonald et al., 2003; see Driver and Noesselt, 2008, and Macaluso and Driver, 2005, for a more extensive overview). The present results indicate that bimodal frequency tagging may amply compensate for the loss of temporal resolution by providing great sensitivity and power to detect tonic and subtle changes in regionally specific background neural activity related to crossmodal influences. Frequency tagging and evoked potentials thus seem to represent somewhat complementary techniques, which probably can be used profitably as such in EEG/MEG research on intermodal attention and crossmodal interactions.

References

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P.-E., Giard, M.H., Bertrand, O., 2007. Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. The Journal of Neuroscience 27, 9252–9261.

Bonnel, A.M., Hafter, E.R., 1998. Divided attention between simultaneous auditory and visual signals. Perception and Psychophysics 60, 179–190.

Bonnel, A.M., Stein, J.F., Bertucci, P., 1992. Does attention modulate the perception of luminance changes? Quarterly Journal of Experimental Psychology A 44, 601–626.

Bohorquez, J., Ozdamar, O., 2008. Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution. Clinical Neurophysiology 119, 2598–2607.

Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vision 10, 433–436.

Brawn, P.T., Snowden, R.J., 2000. Attention to overlapping objects: detection and discrimination of luminance changes. Journal of Experimental Psychology: Human Perception and Performance 26, 342–358.

Deco, G., Zihl, J., 2006. The neurodynamics of visual search. Visual Cognition 14, 1006–1024.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Annual Review Neuroscience 18, 193–222.

Ding, J., Sperling, G., Srinivasan, R., 2006. Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency. Cerebral Cortex 16, 1016–1029.

Downing, C.J., 1988. Expectancy and visual-spatial attention: effects on perceptual quality. Journal of Experimental Psychology Human Perception and Performance 14, 188–202.

Driver, J., Noesselt, T., 2008. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron 57, 11–23.

Duncan, J., 1980. The locus of interference in the perception of simultaneous stimuli. Psychological Review 87, 272–300.

Duncan, J., 1981. Directing attention in the visual field. Perception and Psychophysics 30, 90–93.

Eimer, M., Schroger, E., 1998. ERP effects of intermodal attention and cross-modal links in spatial attention. Psychophysiology 35, 313–327.

Fell, J., Fernndez, G., Klaver, P., Elger, C.E., Fries, P., 2003. Is synchronized neuronal gamma activity relevant for selective attention? Brain Research: Brain Research Review 42, 265–272.

Galambos, R., Makeig, S., Talmacho, P.J., 1981. A 40-Hz auditory potential recorded from the human scalp. Proceedings of the National Academy of Sciences U.S.A. 78, 2643–2647.

Gander, P.E., Bosnyak, R., Wolek, R., Roberts, L.E., 2007. Modulation of the 40-Hz auditory steady-state response by attention during acoustic training. International congress series 1300, 37–40.

Ghazanfar, A.A., Schroeder, C.E., 2006. Is neocortex essentially multisensory? Trends in Cognitive Sciences 10, 278–285.

Giard, M.H., Fort, A., Mouchetant-Rostaing, Y., Pernier, J., 2000. Neurophysiological mechanisms of auditory selective attention in humans. Frontiers in Bioscience 5, D84–D94.

Giard, M.H., Peronnet, F., 1999. Auditory–visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. Journal of Cognitive Neuroscience 11, 473–490.

Gladwin, T.E., Lindsen, J.P., de Jong, R., 2006. Pre-stimulus EEG effects related to response speed, task switching and upcoming response hand. Biological Psychology 72, 15–34.

Gratton, G., Coles, M.G., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology 55, 468–484.

Gutschalk, A., Mase, R., Roth, R., Ille, N., Rupp, A., Hahnel, S., Picton, T.W., Scherg, M., 1999. Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. Clinical Neurophysiology 110, 856–868.

Hawkins, H.L., Hillyard, S.A., Luck, S.J., Mouloua, M., Downing, C.J., Woodward, D.P., 1990. Visual attention modulates signal detectability. Journal of Experimental Psychology: Human Perception and Performance 16, 802–811.

Herrmann, C.S., 2001. Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. Experimental Brain Research 137, 346–353.

Hillyard, S.A., Teder-Salejarvi, W.A., Munte, T.F., 1998. Temporal dynamics of early perceptual processing. Current Opinion in Neurobiology 8, 202–210.

Hoffman, J.E., 1979. A two-stage model of visual search. Perception and Psychophysics 25, 319–327.

Hoffman, J.E., Houck, M.R., MacMillan, F.W., Simons, R.F., Oatman, L.C., 1985. Event-related potentials elicited by automatic targets: a dual-task analysis. Journal of Experimental Psychology: Human Perception and Performance 11, 50–61.

Isreal, J.B., Chesney, G.L., Wickens, C.D., Donchin, E., 1980. P300 and tracking difficulty: evidence for multiple resources in dual-task performance. Psychophysiology 17, 259–273.

Johnson, J.A., Zatorre, R.J., 2005. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. Cerebral Cortex 15, 1609–1620.

Johnson, J.A., Zatorre, R.J., 2006. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. Neuroimage 31, 1673–1681.

Johnson, R., 1984. P300: A model of the variables controlling its amplitude. Annals New York Academy of Sciences 425, 223–229.

Johnson, R., 1986. A triarchic model of p300 amplitude. Psychophysiology 23, 367–384.

Kastner, S., Weerd, P.D., Desimone, R., Ungerleider, L.G., 1998. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. Science 282, 108–111.

Kawahara, J., Lollo, V.D., Enns, J.T., 2001. Attentional requirements is visual detection and identification: evidence from the attentional blink. Journal of Experimental Psychology: Human Perception and Performance 27, 969–984.

Kim, Y.J., Grabowecky, M., Paller, K.A., Muthu, K., Suzuki, S., 2007. Attention induces synchronization-based response gain in steady-state visual evoked potentials. Nature Neuroscience 10, 117125.

Lappin, J.S., Uttal, W.R., 1976. Does prior knowledge facilitate the detection of visual targets in random noise? Perception and Psychophysics 20, 367–374.

Linden, R.D., Picton, T.W., Hamel, G., Campbell, K.B., 1987. Human auditory steady-state evoked potentials during selective attention. Electroencephalography and Clinical Neurophysiology 66, 145–159.

Macaluso, E., Driver, J., 2005. Multisensory spatial interactions: a window onto functional integration in the human brain. Trends in Neurosciences 28, 264–271.

Macmillan, N.A., Creelman, C.D., 2005. Detection Theory: A User's Guide, 2nd ed. Lawrence Erlbaum Associates, Mahwah, N.J.

Malinowski, P., Fuchs, S., Muller, M.M., 2007. Sustained division of spatial attention to multiple locations within one hemifield. Neuroscience Letters 414, 65–70.

Mangun, G.R., Hillyard, S.A., 1990. Allocation of visual attention to spatial locations: tradeoff functions for event-related brain potentials and detection performance. Perception and Psychophysics 47, 532–550.

McCarthy, G., Wood, C.C., 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroencephalography Clinical Neurophysiology 62, 203–208.

McDonald, J.J., Teder-Salejarvi, W.A., Russo, F.D., Hillyard, S.A., 2003. Neural substrates of perceptual enhancement by cross-modal spatial attention. Journal of Cognitive Neuroscience 15, 10–19.

Morgan, S.T., Hansen, J.C., Hillyard, S.A., 1996. Selective attention to stimulus location modulates the steady-state visual evoked potential. Proceedings of the National Academy of Sciences U.S.A. 93, 4770–4774.

Muller, H.J., Humphreys, G.W., 1991. Luminance-increment detection: capacity-limited or not? Journal of Experimental Psychology: Human Perception and Performance 17, 107–124.

Muller, M.M., Hillyard, S., 2000. Concurrent recording of steady-state and transient event-related potentials as indices of visual–spatial selective attention. Clinical Neurophysiology 111, 1544–1552.

Muller, M.M., Picton, T.W., Valdes-Sosa, P., Riera, J., Teder-Salejarvi, W.A., Hillyard, S.A., 1998a. Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28Hz range. Brain Research: Cognitive Brain Research 6, 249–261.

Muller, M.M., Teder-Salejarvi, W., Hillyard, S.A., 1998b. The time course of cortical facilitation during cued shifts of spatial attention. Nature Neuroscience 1, 631–634.

Navon, D., 1985. Attention and performance XI. In: Posner, I., Marin, O.S.M. (Eds.), Erlbaum, Hillsdale, NJ, pp. 133–146.

Navon, D., Miller, J., 2002. Queuing or sharing? a critical evaluation of the single-bottleneck notion. Cognitive Psychology 44, 193–251.

Norman, D., Bobrow, D., 1975. On data-limited and resource-limited processing. Journal of Cognitive Psychology 7, 44–60.

Pastor, M.A., Valencia, M., Artieda, J., Alegre, M., Masdeu, J.C., 2007. Topography of cortical activation differs for fundamental and harmonic frequencies of the steady-state visual-evoked responses. An EEG and PET H_2O^{15} study. Cerebral Cortex 17, 1899–1905.

Pei, F., Pettet, M.W., Norcia, A.M., 2002. Neural correlates of object-based attention. Journal of Vision 2, 588–596.

Pelli, D.G., 1997. The videotoolbox software for visual psychophysics: transforming numbers into movies. Spatial Vision 10, 437–442.

Raymond, J.E., Shapiro, K.L., Arnell, K.M., 1992. Temporary suppression of visual processing in an RSVP task: an attentional blink? Journal of Experimental Psychology Human Perception and Performance 18, 849–860.

Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier, Amsterdam.

Reyes, S.A., Lockwood, A.H., Salvi, R.J., Coad, M.L., Wack, D.S., Burkard, R.F., 2005. Mapping the 40-Hz auditory steady-state response using current density reconstructions. Hearing Research 204, 1–15.

Ross, B., Herdman, A.T., Pantev, C., 2005a. Right hemispheric laterality of human 40 Hz auditory steady-state responses. Cerebral Cortex 15, 2029–2039.

Ross, B., Herdman, A.T., Pantev, C., 2005b. Stimulus induced desynchronization of human auditory 40-Hz steady-state responses. Journal of Neurophysiology 94, 4082–4093.

Ross, B., Herdman, A.T., Wollbrink, A., Pantev, C., 2004a. Auditory cortex responses to the transition from monophonic to pseudo-stereo sound. Neurology and Clinical Neurophysiology 2004, 18.

Ross, B., Picton, T.W., Herdman, A.T., Pantev, C., 2004b. The effect of attention on the auditory steady-state response. Neurology and Clinical Neurophysiology 2004, 22.

Ross, B., Picton, T.W., Pantev, C., 2002. Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. Hearing Research 165, 68–84.

Russo, F.D., Martinez, A., Hillyard, S.A., 2003. Source analysis of event-related cortical activity during visuo-spatial attention. Cerebral Cortex 13, 486–499.

Russo, F.D., Pitzalis, S., Aprile, T., Spitoni, G., Patria, F., Stella, A., Spinelli, D., Hillyard, S.A., 2007. Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. Human Brain Mapping 28, 323–334.

Saupe, K., Widmann, A., Bendixen, A., Muller, M.M., Schroger, E., 2009. Effects of intermodal attention on the auditory steady-state response and the event-related potential. Psychophysiology 46 (2), 321–327 2009.

Sharbrough, F., Chatrian, G., Lesser, R., Luders, H., Nuwer, M., Picton, T., 1991. American Electroencephalographic Society guidelines for standard electrode position nomenclature. Journal of Clinical Neurophysiology 8, 200–202.

Shomstein, S., Yantis, S., 2004. Control of attention shifts between vision and audition in human cortex. Journal of Neuroscience 24, 10702–10706.

Simpson, M.I.G., Hadjipapas, A., Barnes, G.R., Furlong, P.L., Witton, C., 2005. Imaging the dynamics of the auditory steady-state evoked response. Neuroscience Letters 385, 195–197.

Skosnik, P.D., Krishnan, G.P., O'Donnell, B.F., 2007. The effect of selective attention on the gamma-band auditory steady-state response. Neuroscience Letters 420, 223–228.

Sperling, G., Melcher, M.J., 1978. The attention operating characteristic: examples form visual search. Science 202, 315–318.

Stein, B.E., Stanford, T.R., 2008. Multisensory integration: current issues from the perspective of the single neuron. Nature Review Neuroscience 9, 255–266.

Tallon-Baudry, Bertrand, 1999. Oscillatory gamma activity in humans and its role in object representation. Trends in Cognitive Sciences 3, 151–162.

Talsma, D., Doty, T.J., Strowd, R.E., Woldorff, M.G., 2006. Attentional capacity for processing concurrent stimuli is larger across sensory modalities than within a modality. Psychophysiology 43, 541–549.

Talsma, D., Doty, T.J., Woldorff, M.G., 2007. Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? Cerebral Cortex 17, 679–690.

Talsma, D., Kok, A., 2001. Nonspatial intermodal selective attention is mediated by sensory brain areas: evidence from event-related potentials. Psychophysiology 38, 736–751.

Talsma, D., Woldorff, M.G., 2005. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. Journal of Cognitive Neuroscience 17, 1098–1114.

Teder-Salejarvi, W.A., McDonald, J.J., Russo, F.D., Hillyard, S.A., 2002. An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. Brain Research: Cognitive Brain Research 14, 106–114.

Toffanin, P., de Jong, R., Johnson, A., Martens, S., 2009. Using frequency tagging to quantify attentional deployment in a visual divided attention task. International Journal of Psychophysiology 72, 289–298.

Tootell, R.B., Hadjikhani, N.K., Vandu_el, W., Liu, A.K., Mendola, J.D., Sereno, M.I., et al., 1998. Functional analysis of primary visual cortex (V1) in humans. Proceedings of the National Academy of Sciences U.S.A. 95, 811–817.

Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., Theeuwes, J., 2008. Pip and pop: non-spatial auditory signals improve spatial visual search. Journal of Experimental Psychology: Human Perception and Performance 34, 1053–1065.

Wickens, C., Kramer, A., Vanasse, L., Donchin, E., 1983. Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. Science 221, 1080–1082.

Wu, C.T., Weissman, D.H., Roberts, K.C., Woldorff, M.G., 2007. The neural circuitry of the executive control of auditory spatial attention. Brain Research 1134, 187–198.