# Mechanisms of Percept–Percept and Image–Percept Integration in Vision: Behavioral and Electrophysiological Evidence

## Silvia Dalvit and Martin Eimer Birkbeck College, University of London

Previous research has shown that the detection of a visual target can be guided not only by the temporal integration of two percepts, but also by integrating a percept and an image held in working memory. Behavioral and event-related brain potential (ERP) measures were obtained in a target detection task that required temporal integration of 2 successively presented stimuli in the left or right hemifield. Task performance was good when both displays followed each other immediately (percept–percept integration) and when displays were separated by a 300- or 900-ms interval (image–percept integration), but was poor with intermediate interstimulus intervals. An enhanced posterior negativity at electrodes contralateral to the side of the target was observed for percept–percept and for image–percept integration, demonstrating that both are based on spatiotopic representations. However, this contralateral negativity emerged later and was more sustained on trials with long interstimulus intervals, indicating that image–percept integration is slower and involves a sustained activation of working memory.

Keywords: selective attention, working memory, perception, imagery, event-related brain potentials

The distinction between representations that are generated as an immediate result of perceptual objects affecting the visual system (percepts) and representations that are activated in the absence of current visual input (images) is central to models of visual cognition. It has been shown that visual percepts and visual images share functional properties and are implemented by overlapping neural structures (e.g., Awh & Jonides, 2001; Farah, 1988; Kosslyn, 1987). For example, the account of visual imagery proposed by Kosslyn and colleagues assumes that visual images are analog working memory representations that are organized in a spatiotopic fashion, similar to visual percepts (Kosslyn, 1987, 1994; Kosslyn & Thompson, 2003). Evidence for the similarity of percepts and images comes from studies of visual imagery demonstrating that both types of visual representations encode spatial properties of visual objects (e.g., Kosslyn, 1978; Pinker, 1980). Neuropsychological observations have shown that deficits in visual perception and visual imagery are frequently linked (e.g., Farah, 1988), and functional neuroimaging studies have demonstrated that visual brain regions, including occipitoparietal and occipitotemporal areas, are involved in both perception and imagery (Klein et al., 2004; Kosslyn et al., 1999). In addition, experiments comparing visual perception and visual working memory have suggested that both are controlled in a similar fashion by

1

spatial attention (e.g., Awh, Anllo-Vento, & Hillyard, 2000; Nobre et al., 2004). Even though there is now strong evidence for the analog and spatiotopic nature of visual images, alternative accounts of visual imagery are still discussed, such as the hypothesis that images are propositional representations of visual objects (e.g., Pylyshyn, 2002).

If visual percepts and visual images are organized in a similar fashion and are implemented by overlapping brain regions, it should be possible to generate visual representations that combine the information contained in percepts and images and to use these combined representations to guide the detection and discrimination of task-relevant visual events. In a pioneering series of experiments, Brockmole, Wang, and Irwin (2002) investigated whether and how visual images in working memory can be integrated with visual percepts. In these experiments, participants had to combine information from two successively presented grid displays. Both displays contained dots at different grid positions, such that all but one of these positions were filled either in the first or in the second display. The task was to localize the one remaining empty grid position. The critical variable was the interstimulus interval (ISI) between the two displays, which was varied across experimental blocks. Participants performed well when the two displays immediately succeeded each other (0-ms ISI), and also when the ISI was 1,300 ms or longer. In contrast, performance was very poor with intermediate ISIs.

Brockmole et al. (2002) interpreted these results in terms of percept–percept and image–percept integration. When displays immediately follow each other, activity in visual brain areas triggered by the first display is still present by the time activation related to the second display emerges. Perceptual representations of the two displays can therefore be directly combined (percept– percept integration), resulting in good performance. With long intervals between both displays, a short-term memory representation of the first display is available when the second display

This article was published Online First November 1, 2010.

Silvia Dalvit and Martin Eimer, Department of Psychology, Birkbeck College, University of London, England.

Silvia Dalvit is supported by a PhD studentship from the Medical Research Council, UK. Martin Eimer holds a Royal Society–Wolfson Research Merit Award. We thank Monika Kiss for valuable comments and Federica Meconi for technical assistance.

Correspondence concerning this article should be addressed to Martin Eimer, Department of Psychology, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK. E-mail: m.eimer@bbk.ac.uk

arrives, and can be combined with a perceptual representation of the second display (image-percept integration), resulting in task performance that is comparable to the 0-ms ISI condition. However, for intermediate ISIs, a perceptual representation of the first display is no longer available when the second display is presented, and an adequate working memory representation of this display has not yet been formed. Because neither percept-percept integration nor image-percept integration is available to guide target localization, performance is poor. The observation that an ISI of 1,300 ms was sufficient to restore task performance to the level observed for the 0-ms ISI condition led Brockmole et al. to assume that the formation of a stable working memory image of the first grid display requires about 1,300 ms.

To confirm their assumption that successful target localization with long ISIs was based on an integration of images and percepts, rather than on the creation of a working memory representation of the second display and its subsequent combination with a memory representation of the first display (image-image integration), Brockmole et al. (2002) conducted a second experiment. They used a delayed response procedure to estimate the time demands of integration processes in trials where the two displays followed each other immediately, thus enabling percept-percept integration, and in trials where they were separated by a 2,000-ms interval. The estimated integration time was about 300 ms longer with long ISIs. The fact that this time difference was much smaller than the amount of time needed to create and consolidate a working memory image of the second display (about 1,300 ms) appears to rule out the possibility that task performance with long ISIs was based on image-image integration. However, the observation that target localization was delayed when it was based on image-percept integration relative to percept-percept integration is still remarkable, as it suggests that the speed of these two integration processes differs. Image-percept integration processes may be triggered more slowly than percept-percept integration, or may take longer to be completed, perhaps because they require the active maintenance of an image in visual working memory. Alternatively, the speed differences observed by Brockmole et al. in Experiment 2 may be linked to processes that follow integration. For example, response selection might be slower if image-percept integration yields less precise visual representations than percept-percept integration (see Brockmole et al. for further discussion of this possibility).

In summary, the behavioral results reported by Brockmole et al. (2002) strongly suggest that visual images and visual percepts can be successfully integrated, provided that sufficient time is available to form stable working memory representations of taskrelevant information. The aim of the present study was to combine performance and event-related brain potential (ERP) measures to gain new insights into similarities as well as differences between percept-percept and image-percept integration. One goal was to confirm the behavioral findings of Brockmole et al. with a much simpler temporal integration paradigm. Another goal was to obtain direct electrophysiological evidence for the hypothesis that both percept-percept integration and image-percept integration are based on spatiotopically organized visual representations. A third goal was to investigate whether differences in the speed of target detection with image-percept integration as compared with percept-percept integration reflect differences in the time course

of these integration processes themselves or are generated at later stages that follow integration.

As in the Brockmole et al. (2002) study, participants had to perform a visual discrimination task that required the integration of two successively presented displays (D1 and D2). However, there were only two possible target locations, one in each hemifield. Both displays contained two semicircles that varied in their orientation (see Figure 1). Participants had to determine whether the combination of two successive semicircles on the same side yielded a complete circle. Target circles were present on two thirds of all trials, and were always accompanied by a nontarget on the other side. Each display was presented for 10 ms, and the ISI separating the two displays was manipulated. In the 0-ms ISI condition, D1 and D2 followed each other immediately, which should enable percept-percept integration and good target detection performance. As the present task was much simpler than the task used by Brockmole et al. with respect to its demands on spatial working memory, working memory representations of D1 should become available faster than in the earlier study. Therefore, two long ISI conditions (300 and 900 ms) were included, which were expected to allow target detection to be guided by imagepercept integration. At least for the 900-ms ISI condition, task performance should thus be comparable to performance in the 0-ms ISI condition. To investigate the time course of the drop in target detection performance with intermediate ISIs, we also included trials where the interval separating D1 and D2 was 30, 60, or 90 ms. On most of these trials, neither percept-percept nor image-percept integration should be available to guide the detection of target circles, resulting in poor performance relative to the 0-, 300-, and 900-ms ISI conditions.

In addition to measuring participants' target detection performance for all ISI conditions, electroencephalogram (EEG) were recorded during task execution. To identify electrophysiological correlates of successful percept–percept and image–percept integration, ERPs in response to D2 were computed for trials where the two displays either followed each other immediately (percept– percept integration) or were separated by an ISI of 300 or 900 ms (image–percept integration) and the presence of a target circle was correctly reported. ERP analyses focused on lateralized posterior components that are known to be triggered during spatially selective processing in attention and working memory. In many visual search tasks, an enhanced posterior negativity was observed at electrodes PO7/PO8 contralateral to the side of a target. This N2pc component is assumed to reflect the attentional selection of can-



*Figure 1.* Schematic illustration of the sequence of visual events in a target-present trial, where the combination of two successively presented semicircles (D1 and D2) on the right side resulted in a complete circle. In target-absent trials, no combined circle was present on either side. The interstimulus interval (ISI) between the two displays was 0, 30, 60, 90, 300, or 900 ms.

didate target events among nontargets (e.g., Eimer, 1996; Eimer & Kiss, 2008; Luck & Hillyard, 1994; Woodman & Luck, 1999). In addition, a sustained posterior contralateral negativity (SPCN), also referred to as contralateral delay activity (CDA), has been found in working memory tasks, and has been associated with the spatially selective maintenance of working memory representations (McCollough, Machizawa, & Vogel, 2007; Vogel & Machizawa, 2004) and with the spatially directed access to such representations (Eimer & Kiss, 2010). Experiments that required both the spatial selection of target events and their subsequent maintenance in working memory (e.g., Mazza, Turatto, Umiltá, & Eimer, 2007; Vogel & Machizawa, 2004) have shown that the N2pc precedes the SPCN, although both have similar scalp distributions (but see McCollough et al., 2007).

In the present study, target detection required the successful combination of D1 and D2, either through percept-percept integration (in the 0-ms ISI condition) or image-percept integration (in the 300- and 900-ms ISI conditions). If the resulting combined representations are spatiotopic, the selection of a circle target in the left or right hemifield should be associated with an enhanced negativity at contralateral posterior electrodes that emerges after the presentation of D2. Given the spatiotopic nature of perceptual representations, such a contralateral negativity was expected to be elicited in the 0-ms ISI condition where integration is assumed to take place between percepts. The critical question was whether similar lateralized ERP modulations would also be present when the ISI was 300 or 900 ms. This should be the case if the representations that result from image-percept integration and percept-percept integration are structurally similar in terms of their spatiotopic organization (as suggested by Kosslyn, 1987, and Brockmole et al., 2002). The absence of any contralateral posterior negativity following image-percept integration would cast considerable doubt on this hypothesis.

As the emergence of a target-driven contralateral negativity in response to D2 is contingent on a prior successful integration of D1 and D2, the onset latency of this effect can be used to draw inferences about the time demands of percept-percept and imagepercept integration, respectively. An earlier onset of the contralateral posterior negativity on trials where D1 and D2 follow each other immediately relative to trials with long D1-D2 intervals would suggest that percept-percept integration is faster than image-percept integration. This would be in line with the observation by Brockmole et al. (2002, Experiment 2) that target localization is slower when it depends on image-percept integration. The absence of any such latency differences of lateralized posterior ERP modulations would suggest that this behavioral effect is generated at postintegration stages such as response selection. Finally, if image-percept integration involves the spatially selective activation of integrated representations in visual working memory, a sustained posterior contralateral negativity should be elicited with long D1-D2 intervals, but not in the 0-ms ISI condition, where integration is assumed to take place between percepts, without the involvement of visual working memory.

#### Method

# **Participants**

Fifteen paid volunteers participated in this experiment. The data of three participants were not included in the analysis because of an insufficient number of trials per condition after EEG artifact rejection (see below for rejection criteria). The remaining 12 participants (seven women, five men) were 21-36 years old (mean age = 27.5 years), were right-handed, and had normal or corrected-to-normal vision.

## **Stimuli and Procedure**

The experiment was conducted in a dimly lit, sound-attenuated cabin. Stimuli were presented on a CRT monitor (Samsung SyncMaster 1100 MB) with a 100-Hz refresh rate at a viewing distance of 90 cm. E-Prime software (Psychology Software Tools, Pittsburgh, PA) was used for stimulus presentation and behavioral response collection.

Each trial contained two successive displays that included the outlines of two semicircles (angular size of each semicircle:  $0.96^{\circ} \times 0.48^{\circ}$ ) in the left and right hemifield, at a horizontal distance of 1.74° from a central fixation point. Individual semicircles could cover the lower, upper, left, or right half of a full circle (see Figure 1). Fixation point and semicircles were presented in red (CIE *x/y* values: .600/.342) against a green background (CIE *x/y* values: .288/.500). Stimuli and background were equiluminant (9.20 cd/m<sup>2</sup>).

The first display (D1) was presented for 10 ms and was followed by a blank interval of variable duration (see below). Subsequently, a second display (D2) was shown for 10 ms. The interval between the offset of D2 and the onset of D1 on the next trial was 2,200 ms. Participants' task was to report the presence or absence of a combined circle target, that is, a complete circle resulting from the combination of two nonoverlapping semicircles in D1 and D2. Two response keys were vertically arranged in front of the participants, aligned with their body midline. The upper key was used for target-present responses and the lower key for target-absent responses. Six participants signaled the presence of a target with the right index finger and its absence with the left index finger, and this assignment was reversed for the other six participants.

Two thirds of all trials were target-present trials, with target side (left or right) randomly selected across trials. Targets were always accompanied by a nontarget on the opposite side. Nontargets were created by rotating the semicircle in D2 by  $90^{\circ}$  to the left or to the right relative to the D1 semicircle on the same side, such that their combination did not yield a complete circle and ensuring that two successive semicircles on the same side were never identical. Figure 1 shows a trial with a target on the right side and a nontarget on the left side. On one third of all trials, nontargets were presented on both sides.

The main manipulation concerned the ISI between D1 and D2, with six ISI conditions (0 ms, 30 ms, 60 ms, 90 ms, 300 ms, and 900 ms). Results from a pilot study revealed that the circle detection task was very difficult when all six ISI conditions were presented in a randomized fashion in the same block. In this study, performance in the two long ISI conditions was near chance level, and participants consistently reported their inability to perform the task as instructed. Therefore, the experiment was divided into two successive parts, which contained blocks where the ISI was either short (0–90 ms) or long (90–900 ms). The short ISI part was divided into 30 blocks of 30 trials, resulting in a total of 225 trials for each of the four short ISI conditions (referred to as S0, S30, S60, and S90). The long ISI part was divided into 27 blocks of 25

trials, resulting in a total of 225 trials for each of the three long ISIs conditions (L90, L300, and L900). Each block lasted about 1 min, and participants were allowed to take breaks between blocks. Within each block, target-present and target-absent trials and different ISI conditions were presented in random order. ISIs of 90 ms were included in both parts to investigate whether there were systematic differences in target detection performance in the short and long ISI parts that might be linked to different strategies or attentional engagement levels. Six participants completed the short ISI part prior to the long ISI part, and this order was reversed for the other six participants.

#### EEG Recordings and Data Analysis

EEG was DC-recorded with a BrainAmps DC amplifier (upper cutoff frequency 40 Hz, 500-Hz sampling rating) and Ag-AgCl electrodes mounted on an elastic cap from 23 scalp sites (Fpz, F7, F3, Fz, F4, F8, FC5, FC6, T7, C3, Cz, C4, T8, CP5, CP6, P7, P3, Pz, P4, P8, P07, P08, and Oz, according to the extended international 10-20 system). Horizontal electro-oculogram (HEOG) was recorded bipolarly from the outer canthi of both eyes. An electrode placed on the left earlobe served as reference for the online recording, and EEG was rereferenced offline to the average of the left and right earlobes. Electrode impedances were kept below  $5 k\Omega$ .

The main ERP analyses were focused on the three ISI conditions where target detection performance was expected to be good (S0, L300, and L900), and included only target-present trials where a target was correctly reported. The EEG obtained for these trials was epoched into segments starting 100 ms prior to D1 onset and ending 700 ms after D2 onset. EEG was averaged relative to a 100-ms baseline prior to D1 onset, separately for all three ISI conditions and for trials with targets in the left or right hemifield. Trials with saccades (HEOG exceeding  $\pm$  30  $\mu$ V), eye blinks (Fpz exceeding  $\pm$  60  $\mu$ V) or other artifacts (activity at any other electrode site exceeding  $\pm$  80  $\mu$ V) were excluded from analysis. Averaged HEOG waveforms obtained for each participant in the interval between D1 onset and 500 ms after D2 onset were examined to identify residual eye movements on trials with targets on the left or right side. The residual maximal HEOG deviation did not exceed  $\pm 4\mu V$  for any participant.

To quantify lateralized posterior ERP components triggered in response to D2, we computed mean amplitude values obtained at lateral posterior electrodes PO7 and PO8 (where the N2pc component is maximal) for three successive time intervals (200–250 ms, 250–300 ms, and 300–500 ms relative to D2 onset). Mean amplitudes obtained within each of these time windows were analyzed with repeated measures analyses for the factors ISI (0 ms vs. 300 ms vs. 900 ms) and contralaterality (electrode contralateral vs. ipsilateral to the side of a target). Follow-up analyses were conducted separately for each ISI condition.

Target–nontarget discrimination accuracy was analyzed on the basis of *d'* scores obtained for all ISI conditions (see Macmillan & Creelman, 1991, for details). Reaction times (RTs) for correct responses on target-present trials were analyzed for the three ISI conditions where target detection performance was good (S0, L300, and L900) because the percentage of correctly detected targets was generally low in the other ISI conditions (see Table 1), and several participants had very few trials with correct responses

## Table 1

Percentage of Correct Responses on Target-Absent and Target-Present Trials, Shown Separately for All Interstimulus Interval (ISI) Conditions

ISI	Target-absent trials	Target-present trials
SO	89	78
S30	90	38
S60	88	25
S90	87	29
L90	73	50
L300	92	67
L900	96	68

*Note.* S = short ISI conditions of 0, 30, 60, and 90 ms; L = long ISI conditions of 90, 300, and 900 ms.

in some of these conditions. Greenhouse–Geisser corrections for nonsphericity were applied to analyses where appropriate.

#### Results

## **Behavioral Data**

d'. Table 1 shows the percentage of trials where participants correctly reported the presence or absence of a target, separately for each ISI condition. These data were used to compute d' scores, which are shown in Figure 2. Target–nontarget discrimination performance was good in condition S0, dropped dramatically when ISI was increased, but recovered for the two longest ISI conditions (L300 and L900).

A main effect of ISI on d' scores, F(6, 66) = 25.0, p < .001,  $\eta_p^2 = .70$ , confirmed the reliability of this pattern. Further analysis (Bonferroni-adjusted pairwise comparisons) revealed that d'scores observed in conditions S0, L300, and L900 did not differ reliably from each other, but were higher than the d' scores obtained in all other ISI conditions (ps < .05). The d' scores related to conditions S30, S60, S90, and L90 did not differ significantly from each other, except for the score for S30, which was higher than the score for S60 (p < .05). Even though d' did not differ between conditions S90 and L90, there was a difference in response bias, as participants were generally more inclined to report the presence of a circle target in condition L90, which was presented among trials with longer D1-D2 intervals where target detection accuracy was high, than in condition S90 (response criterion c for conditions L90 and S90, based on mean accuracy on target-present and target-absent trials: 0.31 and 0.84, respectively, t[11] = 2.39, p < .05).

**Reaction times.** RTs on trials where the presence of a target was correctly reported differed reliably between the three conditions where target detection performance was good (S0, L300, and L900). Mean RTs in condition S0 (612 ms) were faster than in conditions L300 and L900 (684 and 660 ms, respectively), resulting in a main effect of ISI, F(2, 22) = 9.9, p < .05,  $\eta_p^2 = .47$ . Follow-up analyses confirmed significant RT differences between conditions S0 and conditions L300 and L900 (ps < .05), whereas RTs in conditions L300 and L900 did not differ reliably. For intermediate ISIs, mean RTs on trials where targets were reported



*Figure 2. d'* scores obtained for all interstimulus interval (ISI) conditions, shown separately for blocks with short ISIs (S0, S30, S60, and S90, solid line), and for blocks with long ISIs (L90, L300, and L900, dashed line). Error bars represent standard errors of the mean.

correctly were 709 ms, 693 ms, 699 ms, and 778 ms for conditions S30, S60, S90, and L90, respectively.

## **ERP** Data

Figure 3 shows grand-averaged ERPs obtained at electrodes PO7/PO8 ipsilateral and contralateral to the side of a circle target. ERPs are time-locked to D1 onset, and are shown separately for the three ISI conditions S0, L300, and L900. A pronounced posterior contralateral negativity was triggered in all three ISI conditions. This effect appeared to start earlier in the S0 condition relative to the L300 and L900 conditions, but was more sustained in the two long ISI conditions. This can be seen more clearly in Figure 4, which presents difference waveforms obtained by subtracting ERPs at ipsilateral electrodes from contralateral ERPs, separately for the three ISI conditions. Topographic maps of the scalp distribution of these contralateral–ipsilateral differences for these three conditions are shown in Figure 3 (top left panels).

These observations were confirmed by statistical analyses. In the 200- to 250-ms time interval after D2 onset, a main effect of contralaterality, F(1, 11) = 28.69, p < .001,  $\eta_p^2 = .72$ , was accompanied by an interaction between ISI and contralaterality,  $F(2, 22) = 6.79, p < .05, \eta_p^2 = .38$ . Follow-up analyses conducted separately for each ISI condition revealed significant differences between contralateral and ipsilateral ERPs for condition S0, t(11) = 4.07, p < .01, but not for conditions L300 and L900, t(11) = 1.9, p = .08, and t(11) = 1.74, p = .11, respectively,indicating that a posterior contralateral negativity emerged earlier in the S0 condition than in the L300 and L900 conditions. In the subsequent 250- to 300-ms time interval after D2 onset, a main effect of contralaterality was again observed, F(1, 11) = 125.27, p < .001,  $\eta_p^2 = .92$ , but no ISI  $\times$  contralaterality interaction was present, F(2, 22) = 0.69, p = .513,  $\eta_p^2 = .06$ , indicating that during this time window, a posterior contralateral negativity of similar size was elicited in all three ISI conditions. In the 300- to 500-ms time interval after D2 onset, analyses revealed a main effect of contralaterality,  $F(1, 11) = 28.26, p < .001, \eta_p^2 = .72$ , as well as an interaction between ISI and contralaterality, F(2, 22) = 4.85, p < .05,  $\eta_p^2 = .31$ . Follow-up analyses for each ISI conditions demonstrated that, in marked contrast to the 200- to 250-ms time window, significant differences between contralateral and ipsilateral ERPs were now present in the two longer ISI conditions, t(11) = 4.11, p < .01, and t(11) = 5.11, p < .001, for conditions L300 and L900, respectively, but not for condition S0, t(1,11) = 1.53, p = .15, thus demonstrating that the posterior contralateral negativity was more sustained when the D1–D2 interval was longer.

To demonstrate more directly that the posterior contralateral negativity emerged earlier in condition S0 relative to conditions L300 and L900, we determined the onset of this effect by jackknife-based analyses (Miller, Patterson, & Ulrich, 1998; Ulrich & Miller, 2001) that were based on individual contralateral-ipsilateral difference waveforms obtained for these three conditions. The onset of the contralateral negativity was defined as the post-D2 latency where difference waveforms reached 25% of their maximal amplitude. Estimated onset latencies for conditions S0, L300, and L900 were 196 ms, 226 ms, and 237 ms after D2 onset, respectively, and these estimates correspond well with the effect onsets suggested by Figure 4. Statistical analyses of these onset estimates, where F and t values were corrected according to the formula described by Ulrich and Miller (2001), confirmed that the onset latency of the posterior contralateral negativity was reliably affected by ISI,  $F_c(2, 22) = 14.3$ , p <.01,  $\eta_p^2 = .57$ . It was significantly earlier in condition S0 than in conditions L300 and L900,  $t_c(11) = 2.97$ , p < .05, and  $t_c(11) =$ 10.26, p < .001, one-tailed, respectively. In contrast, onset latency estimates for the two long ISI conditions did not differ reliably,  $t_{\rm c}(11) = 1.34, p > .05.$ 

As can be seen in Figure 4, the posterior contralateral negativity in condition S0 that emerged about 200 ms after D2 onset was preceded by a brief period where a lateralized posterior activity of opposite polarity was present. A post hoc analysis confirmed that ERPs were more positive at contralateral as compared with ipsi-



*Figure 3.* Grand-averaged event-related potentials (ERPs) elicited in the interval between D1 onset and 700 ms after D2 onset at electrodes PO7/PO8 contralateral and ipsilateral to the side of a circle target on trials where targets were correctly reported. ERPs are shown separately for conditions S0, L300, and L900 (S = short; L = long), relative to a 100-ms baseline prior to D1. All statistical analyses were performed on ERPs elicited after D2 onset (indicated by vertical dashed lines). Top left panels show scalp distribution maps of the difference between contralateral and ipsilateral ERPs during the 200- to 250-ms interval after D2 onset (S0 interstimulus interval [ISI] condition) or during the 300- to 500-ms post-D2 interval (L300 and L900 ISI conditions). These maps were constructed by spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989) after mirroring the ipsilateral–contralateral difference waveforms to obtain symmetrical voltage values for both hemispheres.

lateral electrodes in the 150- to 180-ms time interval after D2 onset, t(11) = 2.97, p < .05, in condition S0. This unexpected effect, which was not present in the L300 and L900 conditions (see Figure 4), is likely to reflect a sensory asymmetry between the visual hemifields containing the target and the nontarget that is specific to condition S0 where D2 followed D1 immediately (see Discussion for more details).

Figure 5 shows grand-averaged ERPs obtained at PO7/PO8 ipsilateral and contralateral to the circle target side, for conditions S90 and L90 where target detection performance was poor (see Figure 2), collapsed across trials with correct and incorrect responses. In contrast to the S0, L300, and L900 conditions, no systematic contralateral negativity appears to have been elicited in these trials. This was confirmed by paired t tests that compared



*Figure 4.* Difference waves obtained in the 700-ms interval after D2 onset by subtracting ipsilateral from contralateral event-related potentials (ERPs), shown separately for interstimulus interval (ISI) conditions S0 (dashed black line), L300 (solid gray line), and L900 (solid black line; S = short; L = long). Note that these difference waveforms were computed relative to a 100-ms baseline prior to D1 onset.

contralateral and ipsilateral ERP amplitudes obtained in the S90 and L90 conditions for the 200- to 250-ms, 250- to 300-ms, and 300- to 500-ms intervals after D2 onset. No reliable contralateral negativity was observed in any of these time intervals, and there were also no differences of nonlateralized posterior ERPs between these two conditions in any of these time windows (ps > .05).

## Discussion

The aim of this study was to combine behavioral and electrophysiological measures to gain new insights into the processes that underlie the integration of perceptual representations and representations stored in working memory. A simplified version of the temporal integration paradigm used by Brockmole et al. (2002) was used to study the time course of percept-percept and imagepercept integration, as well as the properties of the combined representations that result from these two types of integration.

The behavioral results fully confirmed the observations of Brockmole et al. (2002). As predicted, target detection performance was strongly affected by manipulating the time interval between the two to-be-integrated displays, and this modulation followed a U-shaped function (see Figure 2). Performance was good when the two displays followed each other immediately and also when they were separated by longer ISIs (300 or 900 ms). In contrast, performance was poor with intermediate ISIs. In line with the interpretation proposed by Brockmole et al., this pattern of results suggests that target localization in the 0-ms ISI condition was guided by percept-percept integration as the activation of visual areas triggered by D1 was still present when the second display arrived. With long ISIs, a stable working memory representation of D1 was available by the time that the second display was presented, and target localization was based on image-percept integration. It is notable that performance in the present study was equally good with ISIs of 300 and 900 ms, suggesting that a 300-ms interval was sufficient to form an adequate working memory representation of the first display. The fact that generating such representations took substantially longer (about 1,300 ms) in the Brockmole et al. study is most likely due to the fact that the number of to-be-integrated items was much higher than in the current experiment (see also Vogel, Woodman, & Luck, 2006, for evidence that the time required to consolidate items in visual working memory increases as memory set size increases).

The observation that target detection was poor with ISIs of 30, 60, and 90 ms mirrors the finding of Brockmole et al. (2002) that task performance deteriorated for intermediate ISIs. The drop in performance for D1–D2 intervals between 30 and 90 ms is consistent with the assumption that neither perceptual nor working memory representations of D1 were available for percept–percept or image–percept integration, respectively: Perceptual representations of D1 are subject to rapid decay, and the formation of working memory representations of D1 cannot be completed within 100 ms after D1 presentation. The performance drop for



*Figure 5.* Grand-averaged event-related potentials (ERPs) elicited in the interval between D1 onset and 700 ms after D2 onset at electrodes PO7/ PO8 contralateral and ipsilateral to the side of a circle target in conditions S90 and L90 (S = short; L = long), relative to a 100-ms baseline prior to D1. Waveforms are collapsed across trials with correct and incorrect responses, and D2 onset is indicated by vertical dashed lines.

intermediate ISIs is unlikely to be primarily due to low-level backward masking, as circle targets resulted from the combination of two successively presented nonoverlapping objects. Thus, the absence of adequate perceptual or working memory representations of D1 provides the most plausible explanation for the poor target detection performance with intermediate ISIs.

Target detection performance was slightly better in condition S30 than in condition S60, suggesting that on some trials, visual activation triggered by D1 persisted for several tens of milliseconds and was thus available for percept-percept integration. It should also be noted that even though performance in the S60, S90, and L90 conditions was poor, it was still above chance, which may be due to the occasional persistence of residual visual activity for up to 100 ms. Participants adopted a more liberal response criterion in condition L90 relative to condition S90, presumably because the L90 condition was presented among trials with longer D1-D2 intervals that facilitated target detection. In contrast, the S90 condition was paired with condition S0 and with two other ISI conditions (S60 and S30) that were associated with poor target detection. However, d' scores did not differ between conditions L90 and S90, demonstrating that such context-dependent differences in response bias had no impact on the efficiency of targetnontarget discrimination.

A final notable behavioral result was that target detection was about 50 ms faster in trials where the two displays followed each other immediately relative to trials where they were separated by intervals of 300 or 900 ms. This suggests that image–percept integration is slower than percept–percept integration, in line with previous observations by Brockmole et al. (2002, Experiment 2), who found a 300-ms difference in target localization speed between these two types of integration processes. The fact that this difference was much smaller in the present study is most likely because of a simplified task design.

In summary, the pattern of behavioral results obtained in the present study fully confirmed previous findings by Brockmole et al. (2002), in spite of the fact that a much simpler temporal integration task was employed here. They are in line with the hypothesis that target detection in the 0-ms ISI condition was based on percept-percept integration, whereas image-percept integration was operative with ISIs of 300 and 900 ms. If this hypothesis is correct, the ERP data obtained during the performance of this task can be used to provide new insights into the nature of the combined representations that result from percept-percept and image-percept integration and into the time course of these integration processes themselves.

The pattern of ERP results obtained for the 0-ms ISI condition was in line with predictions. An enhanced negativity was triggered at posterior electrodes PO7/PO8 contralateral to the side of successfully detected targets. This contralateral negativity started 200 ms after D2 onset, and remained present for about 100 ms (see Figures 3 and 4). Similar N2pc components have previously been observed in visual search tasks in response to targets among nontarget distractor items (e.g., Eimer, 1996; Luck & Hillyard, 1994), and were interpreted to reflect the spatially directed attentional selection of such targets. The presence of an N2pc in the 0-ms ISI condition is not surprising, as target detection was assumed to be guided by percept–percept integration, and perceptual representations are known to be spatiotopically organized. As can be seen in Figure 4, the N2pc was preceded by a transient ERP deflection of opposite polarity (a contralateral positivity) in the 0-ms ISI condition. This early modulation of posterior ERPs may reflect a sensory asymmetry between the target and nontarget hemifields. Targets were by definition always successively presented nonoverlapping semicircles, whereas the semicircles in D1 and D2 partially overlapped on the nontarget side (see Figure 1). In the 0-ms ISI condition, where D1 and D2 were each presented for 10 ms in immediate succession, the color of these overlapping parts on the nontarget side was perceived as slightly more saturated than the nonoverlapping contours of a target circle, and this might have triggered a slightly asymmetric early visual response. No such early visual asymmetry was present in the two long ISI conditions, where there was no perceptual overlap between D1 and D2. However, the early positivity in the 0-ms ISI condition was not due to a contralateral enhancement of the P1 component, which is typically most sensitive to low-level visual differences, but instead reflected a small differential modulation of N1 amplitudes at contralateral versus ipsilateral posterior electrodes (see Figure 3, top right panel). It should be noted that in several previous ERP studies (e.g., Eimer & Kiss, 2008), a similar early contralateral positivity was found to precede the N2pc, even though symmetrical visual displays were used, which may suggest that this effect is not exclusively due to a sensory imbalance between visual fields. In any case, the presence of this transient early difference between ISI conditions implies that subsequent differences in lateralized posterior ERPs need to be interpreted with caution. However, the fact that this early ERP modulation was opposite in polarity to the subsequent N2pc makes it unlikely that its presence will have contributed to the earlier onset of this component in the 0-ms ISI condition relative to the two long ISI conditions. If anything, an earlier contralateral positivity should have resulted in a delay of N2pc latencies in this condition.

One central finding of the present study was that a posterior negativity contralateral to the side of a successfully detected circle target was not just triggered in the 0-ms ISI condition, but also on trials where the D1-D2 interval was 300 or 900 ms, and target detection was assumed to be guided by image-percept integration. Previous ERP research has demonstrated that representations in visual working memory are spatiotopically organized (e.g., Dell'Acqua, Sessa, Toffanin, Luria, & Jolicoeur, 2010; Eimer & Kiss, 2010; Kuo, Rao, Lepsien, & Nobre, 2009; McCollough, Machizawa, & Vogel, 2007; Vogel & Machizawa, 2004). The presence of a pronounced contralateral negativity in the 300- and 900-ms ISI conditions extends these observations by suggesting that this also applies to representations that result from an integration of percepts and images. This is in line with Kosslyn's model of visual mental imagery, which includes a spatiotopic visual buffer that receives input from both perception and memory. Once information has entered this buffer, it is processed in the same fashion, regardless of the source of the original input (see Kosslyn, 1994, for more details).

In contrast to the presence of reliable posterior negativities contralateral to the side where a target circle was detected in the 0-ms, 300-ms, and 900-ms ISI conditions, no such lateralized effects were observed when the D1–D2 ISI was 90 ms (S90 and L90; see Figure 5). This difference is fully in line with the observation that task performance was poor when D1 and D2 were separated by a 90-ms interval. If neither percept–percept nor image–percept integration is available to guide target detection,

lateralized posterior ERP components should be strongly reduced or absent. This prediction was confirmed, providing further support for the hypothesis that these components are elicited when such integration processes are successful.

Even though a contralateral posterior negativity was elicited when D2 followed D1 immediately and when the two displays were separated by a 300- or 900-ms interval, there were also important differences in the time course of this ERP effect between ISI conditions. Its onset was delayed about 40 ms with long ISIs (see Figure 4), which strongly suggests that the detection of target circles was slower when it was based on image-percept integration than when it was guided by percept-percept integration. It is interesting that this onset delay between the 0-ms and 900-ms ISI conditions was almost identical to the RT difference between these conditions (41 ms vs. 48 ms). Brockmole et al. (2002) discussed two possible sources for their observation that image-percept integration was slower than percept-percept integration, an increase in the duration of the integration process itself that may be linked to additional processing demands associated with maintaining working memory representations, or a delay of later response selection stages. The observation that the RT delay for target detection based on image-percept integration with a 900-ms ISIs as compared with percept-percept integration can be almost completely accounted for by the onset difference of a lateralized posterior ERP marker of spatially specific target selection is clearly inconsistent with the hypothesis that this RT difference is primarily generated at a late response-related stage. If imagepercept integration requires more time than percept-percept integration, access to representations that result from image-percept integration should be delayed relative to the access to purely perceptual representations. This prediction is in line with the delayed onset latency of contralateral posterior negativities in the two long ISI conditions relative to the 0-ms ISI condition. However, given that these ERP effects reflect the spatially selective access to representations resulting from percept-percept or imagepercept integration and not the integration process itself, the alternative possibility remains that it is this access rather than the preceding integration that is faster for the 0-ms ISI condition. It should also be noted that the RT difference between the 0-ms and 300-ms ISI conditions (72 ms) was considerably larger than the corresponding N2pc onset latency difference (30 ms). This might suggest that when D1-D2 intervals are relatively short, response latencies are determined not just by the time demands of the image-percept integration process, but also by subsequent decision and response selection processes.

It is unlikely that attentional or strategy differences or different levels of arousal between blocks with short and long ISIs are responsible for the observed differences between the 0-ms ISI condition and the two long ISI conditions. The comparison between the S90 and L90 conditions (which were physically identical, but were delivered in short vs. long ISI blocks) demonstrated that even though participants adopted a more liberal response criterion in the latter condition, perceptual sensitivity did not differ. The direct comparison of lateralized and nonlateralized posterior ERPs obtained in these two conditions revealed no differences (see Figure 5), indicating that general arousal levels were comparable. In addition, the time interval where N2pc onset latency differences between ISI conditions were observed (200–250 ms after D2 onset) appears too early to be affected by differences in response bias.

The contralateral posterior negativity observed with long ISIs was not only delayed, but was also more sustained than the contralateral negativity elicited in the 0-ms ISI condition. In the latter condition, this effect returned to baseline about 300 ms after D2 onset, whereas it remained present during the 300- to 500-ms time window for ISIs of 300 and 900 ms (see Figures 3 and 4). This sustained contralateral negativity observed with long ISIs may reflect the additional spatially selective activation of visual working memory. As noted earlier, a lateralized SPCN (or CDA) component has previously been observed in tasks that require maintenance of or access to representations in working memory (e.g., Eimer & Kiss, 2010; Mazza et al., 2007; Vogel & Machizawa, 2004). The observation that the contralateral negativity was more sustained in conditions where target detection was guided by image-percept integration may indicate a stronger or temporally more extended involvement of spatial working memory than in conditions where targets could be detected by integrating two perceptual representations.

Alternatively, the presence of a more sustained contralateral negativity in the 300- and 900-ms ISI conditions might simply reflect a greater variability in the onset of spatially selective attentional target selection relative to the 0-ms ISI condition, resulting in a temporally smeared N2pc component. This could be the case if the duration of image-percept integration was much more variable across trials than the duration of percept-percept integration. However, such an increase in the latency jitter of an ERP component is typically accompanied by substantial decrease in its amplitude, and there was little evidence that the size of the contralateral posterior negativity was reduced in the two long ISI conditions (see Figure 4). Although this observation makes it unlikely that the sustained contralateral negativity observed with long ISIs is due to an increased latency variability of the N2pc, the hypothesis that it reflects an SPCN component that is associated with the spatially selective activation of working memory will need to be addressed and confirmed in future experiments.

As can be seen in Figure 3, a tendency for ERPs contralateral to the target side to be more negative than ipsilateral ERPs was present in conditions L300 and L900 around 400 ms after D1, well before the onset of any contralateral negativity in response to D2. This may seem puzzling, as the side of the target could be determined only once D2 was presented. Paired t tests performed on ERP mean amplitudes obtained in the 300- to 500-ms interval after D1 onset confirmed reliable differences between contralateral and ipsilateral ERPs for the L300 and L900 conditions, ts(11) > 2.3, ps < .05. These differences may reflect a spontaneous early attentional bias toward one or the other side of the working memory representation of D1. Even though the direction of such a bias will vary randomly across trials, it should increase or decrease the probability of successful target detection following D2 onset. As the ERPs shown in Figure 3 for the L300 and L900 conditions represent those two thirds of all trials where targets were correctly detected, it is likely that a fortuitous allocation of attention to the side of the target occurred more frequently on these trials than on trials where targets remained undetected. In support of this hypothesis, the early contralateral negativity was no longer present when ERPs for conditions L300 and L900 were computed across target-present trials with correct and incorrect responses, ts(11) < 1.

It should be noted that the image-percept integration hypothesis proposed by Brockmole et al. (2002) to account for the effects of ISI in empty-cell detection tasks has not gone unchallenged. As pointed out by Hollingworth, Hyun, and Zhang (2005), the number of dots that was apparently retained in working memory in the long ISI conditions of the Brockmole et al. study substantially exceeds current estimates of working memory capacity (e.g., Luck & Vogel, 1997). This observation led Jiang, Kumar, and Vickery (2005) to propose an alternative interpretation of these findings. According to those authors, the good task performance observed by Brockmole et al. for long D1-D2 ISIs was not based on image-percept integration, but was due to the fact that observers adopted a "negative conversion" strategy: They initially formed a working memory representation of the unfilled grid positions in the first array, which was then compared with the second array to detect the one remaining empty position. In other words, target detection was based on a direct comparison between a working memory and a perceptual representation, analogous to the mechanisms that underlie successful change detection (Rensink, 2002), and not on image-percept integration. This alternative interpretation might also account for the behavioral and ERP effects observed in the present experiment. Instead of integrating of working memory representation of D1 with a perceptual representation of D2 to detect target circles in the long ISI conditions, participants instead may have formed working memory representations of the mirror image of both semicircles contained in D1, which were then compared with D2. This would have resulted in the attentional selection of the matching target semicircle, as reflected by a contralateral posterior negativity, but would not involve an imagepercept integration process as postulated by Brockmole et al.

Several considerations make this alternative interpretation of the current findings unlikely. According to Jiang et al. (2005), participants are forced to adopt the negative conversion strategy in conditions where working memory capacity limitations prevent successful image-percept integration. This may have conceivably been the case in the Brockmole et al. (2002) study, where relatively large grid arrays (4  $\times$  4 or 5  $\times$  5) were employed, but such capacity limitations would not have played a major role in the present study, because each display contained only two items. In addition, generating mirror-image memory representations of D1 that enable the attentional selection of target locations in D2 is a multistage process that requires the perceptual encoding of both D1 items, their mirror rotation, and the formation of a working memory representation of the rotated items. It is plausible to assume that these sequential processes will take several hundred milliseconds (see also Jiang et al. for an analogous argument). If target detection in the long ISI conditions of the present study was guided by this time-consuming strategy, performance should have been much better in the 900-ms ISI condition than in the 300-ms ISI condition. In fact, no such performance difference was observed, and there were also no reliable differences in the ERP correlates of target detection between these two ISI conditions. These considerations suggest that task performance in the long ISI conditions was indeed based on image-percept integration, as described by Brockmole et al., and not on the alternative negative conversion strategy suggested by Jiang et al.

In summary, the present study has demonstrated that the integration of two percepts and the integration of a perceptual and a working memory representation both operate within a spatiotopic reference frame. However, these two integration mechanisms differ in their time demands, such that the spatially selective access to representations that are based on image–percept integration is delayed relative to the access to representations that are generated by percept–percept integration.

#### References

- Awh, E., Anllo-Vento, L., & Hillyard, S. A. (2000). The role of spatial selective attention in working memory for locations: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 12, 840– 847.
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, 5, 119–126.
- Brockmole, J. R., Wang, R. F., & Irwin, D. E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 315–334.
- Dell'Acqua, R., Sessa, P., Toffanin, P., Luria, R., & Jolicoeur, P. (2010). Orienting attention to objects in visual short-term memory. *Neuropsychologia*, 48, 419–428.
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, 99, 225–234.
- Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. *Journal* of Cognitive Neuroscience, 20, 1423–1433.
- Eimer, M., & Kiss, M. (2010). An electrophysiological measure of access to representations in visual working memory. *Psychophysiology*, 47, 197–200.
- Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, 95, 307–317.
- Hollingworth, A., Hyun, J.-S., & Zhang, W. (2005). The role of visual short-term memory in empty cell localization. *Perception & Psychophysics*, 67, 1132–1343.
- Jiang, Y., Kumar, A., & Vickery, T. J. (2005). Integrating sequential arrays in visual short-term memory. *Experimental Psychology*, 52, 39–46.
- Klein, I., Dubois, J., Mangin, J.-F., Kherif, F., Flandin, G., Poline, J.-B., . . . Le Bihan, D. (2004). Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, 22, 26–31.
- Kosslyn, S. M. (1978). Measuring the visual angle of the mind's eye. Cognitive Psychology, 10, 356–389.
- Kosslyn, S. M. (1987). Seeing and imaging in the cerebral hemispheres: A computational approach. *Psychological Review*, 94, 148–175.
- Kosslyn, S. M. (1994). Image and brain: The resolution of the imagery debate. Cambridge: MA: MIT Press.
- Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., . . . Alpert, N. M. (1999, April 2). The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science*, 284, 167–170.
- Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, 129, 723–746.
- Kuo, B. C., Rao, A., Lepsien, J., & Nobre, A. C. (2009). Searching for targets within the spatial layout of visual short-term memory. *Journal of Neuroscience*, 29, 8032–8038.
- Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1000–1014.
- Luck, S. J., & Vogel, E. K. (1997, November 20). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.

- Macmillan, N. A., & Creelman, C. D. (1991). Detection theory: A user's guide. Cambridge, England: Cambridge University Press.
- Mazza, V., Turatto, M., Umiltá, C., & Eimer, M. (2007). Attentional selection and identification of visual objects are reflected by distinct electrophysiological responses. *Experimental Brain Research*, 181, 531– 536.
- McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex*, 43, 77–94.
- Miller, J., Patterson, T., & Ulrich, R. (1998). A jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, 35, 99– 115.
- Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vanderberghe, R., & Mesulam, M. M. (2004). Orienting attention to location in perceptual versus memory representations. *Journal of Cognitive Neuroscience*, 16, 363–373.
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72, 184–187.
- Pinker, S. (1980). Mental imagery and the third dimension. Journal of Experimental Psychology: General, 109, 354–371.

- Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, 25, 157–182.
- Rensink, R. A. (2002). Change detection. Annual Review of Psychology, 53, 255–277.
- Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology*, 38, 816–827.
- Vogel, E. K., & Machizawa, M. G. (2004, April 15). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428, 748–751.
- Vogel, E. K., Woodman, F. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1436–1451.
- Woodman, G. F., & Luck, S. J. (1999, August 26). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, 400, 867–869.

Received May 21, 2009 Revision received March 2, 2010 Accepted March 8, 2010

# **Online First Publication**

APA-published journal articles are now available Online First in the PsycARTICLES database. Electronic versions of journal articles will be accessible prior to the print publication, expediting access to the latest peer-reviewed research.

All PsycARTICLES institutional customers, individual APA PsycNET<sup>®</sup> database package subscribers, and individual journal subscribers may now search these records as an added benefit. Online First Publication (OFP) records can be released within as little as 30 days of acceptance and transfer into production, and are marked to indicate the posting status, allowing researchers to quickly and easily discover the latest literature. OFP articles will be the version of record; the articles have gone through the full production cycle except for assignment to an issue and pagination. After a journal issue's print publication, OFP records will be replaced with the final published article to reflect the final status and bibliographic information.