



BIROn - Birkbeck Institutional Research Online

Berggren, Nick and Eimer, Martin (2018) Object-based target templates guide attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance* 44 (9), pp. 1368-1382. ISSN 0096-1523.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/21065/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

Object-based target templates guide attention during visual search

Nick Berggren* and Martin Eimer

Department of Psychological Sciences, Birkbeck College, University of London,
Malet Street, London WC1E 7HX, UK

* Corresponding author

Phone: 0044 20 76316522

Fax: 0044 20 76316312

Email: nbergg01@mail.bbk.ac.uk

ABSTRACT

During visual search, attention is believed to be controlled in a strictly feature-based fashion, without any guidance by object-based target representations. To challenge this received view, we measured electrophysiological markers of attentional selection (N2pc component) and working memory (SPCN) in search tasks where two possible targets were defined by feature conjunctions (e.g., blue circles and green squares). Critically, some search displays also contained nontargets with two target features (incorrect conjunction objects, e.g., blue squares). Because feature-based guidance cannot distinguish these objects from targets, any selective bias for targets will reflect object-based attentional control. In Experiment 1, where search displays always contained only one object with target-matching features, targets and incorrect conjunction objects elicited identical N2pc and SPCN components, demonstrating that attentional guidance was entirely feature-based. In Experiment 2, where targets and incorrect conjunction objects could appear in the same display, clear evidence for object-based attentional control was found. The target N2pc became larger than the N2pc to incorrect conjunction objects from 250 ms post-stimulus, and only targets elicited SPCN components. This demonstrates that after an initial feature-based guidance phase, object-based templates are activated when they are required to distinguish target and nontarget objects. These templates modulate visual processing and control access to working memory, and their activation may coincide with the start of feature integration processes. Results also suggest that while multiple feature templates can be activated concurrently, only a single object-based target template can guide attention at any given time.

Keywords: selective attention; top-down control; spatial attention; feature-based attention; object-based attention; event-related brain potentials

PUBLIC SIGNIFICANCE STATEMENT

When we look within our environment for an object (e.g., our cell phone), current models of attention assume that our visual system adopts strictly feature-specific mental templates to guide our search (the colour silver, a rectangular shape). Here, we demonstrate that search templates for objects can also guide visual attention. Using electrophysiological recordings, we show that during search for two possible target objects (blue circle, green square), attentional control can distinguish these targets from other objects that contain both target features in an incorrect combination (e.g., blue squares) within a quarter of a second after exposure. These results challenge the presumed architecture of early attentional control by demonstrating that search can be guided rapidly by mental representations of visual objects.

INTRODUCTION

Visual search for specific target objects is guided by observers' knowledge about their properties. In some search tasks, targets are defined by simple features (e.g., colour, size, shape, or orientation), while in others, targets are objects defined by a conjunction of features. Such target-defining features or feature conjunctions are assumed to be represented as attentional templates (e.g., Duncan & Humphreys, 1992) or attentional control sets (e.g., Folk, Remington, & Johnston, 1992). Once a particular target template is activated, stimuli that match this template will attract attention, while stimuli with non-matching features will not (e.g., Folk et al., 1992; Folk & Remington, 1998; Eimer & Kiss, 2008). While the role of attentional templates in the guidance of visual search is undisputed (e.g., Wolfe, 2007), the way in which target features and objects are represented in search templates is not yet fully understood. When targets are defined by a specific feature such as a particular colour, search templates are necessarily feature-specific. However, in search tasks where target objects are defined by a conjunction of features, two different types of target templates are conceivable. On the one hand, templates for feature conjunctions may still be feature-based, operating independently for each target feature. For example, during search for blue circles among red circles and blue squares, a colour template will bias attention towards blue objects and a shape template will produce an attentional bias for circles, but there will not be an additional object-based guidance process that is selectively tuned to blue circles only. Leading models of attentional control such as Guided Search (Wolfe, 1994; 2007) assume that the selection of particular objects during visual search operates in such a feature-based fashion. Once selected, objects are identified by comparing them to long-term memory representations, before they are categorised as targets or nontargets by determining whether they match the current memory set for target objects. This three-stage model has been proposed by Cunningham and Wolfe (2014) as an account of hybrid visual and memory search tasks with multiple possible target objects. In this mode, early visual selection is controlled by features, whereas the subsequent identification and target/nontarget discrimination stages are based on comparisons with stored object representations. An alternative possibility is that object-based target

templates are already involved in the early guidance of attentional selectivity during conjunction search. If the allocation of attention can be controlled by representations where individual target features are bound within a single integrated object file (Kahneman, Treisman, & Gibbs, 1992; Wheeler & Treisman, 2002), attentional biases should emerge for target objects only (e.g., blue circles), and not for objects with only one target-matching feature (e.g., red circles or blue squares). The question addressed in this study is whether the guidance of attention during conjunction search is entirely feature-based, fully object-based, or involves both feature-based and object-based target templates.

Because search templates bias visual processing towards template-matching features and objects, the nature of these templates can be identified by measuring their on-line effects on visual activation. Event-related brain potentials (ERP) offer temporally precise markers of such template-controlled attentional biases. The most prominent ERP marker of attentional selectivity is the N2pc component, which is an enhanced negativity that is triggered at posterior scalp electrodes contralateral to candidate target objects in visual search displays. The N2pc typically emerges between 180-200 ms after search display onset, is generated in ventral extrastriate visual areas (Hopf et al., 2000), and is assumed to reflect the allocation of attention to possible targets (e.g., Luck & Hillyard, 1994; Eimer, 1996; Mazza, Turatto, Umiltà, & Eimer, 2007). The sensitivity of the N2pc to search templates for target features has been demonstrated in spatial cueing tasks where task-irrelevant colour cues presented prior to search displays triggered an N2pc only when they matched the current target colour but not when they had a different non-matching colour (Eimer & Kiss, 2008; Lien, Ruthruff, Goodin, & Remington, 2008; Leblanc, Prime, & Jolicoeur, 2008; Eimer, Kiss, Press, & Sauter, 2009). While the N2pc reflects the allocation of attention to possible target objects during visual search (i.e., the early visual selection stage proposed by Cunningham & Wolfe, 2014), another lateralised ERP component that is often observed in visual search tasks (sustained posterior contralateral negativity/SPCN; Mazza, Turatto, Umiltà, & Eimer, 2007; Jolicoeur, Brisson, & Robitaille, 2008) is linked to the subsequent processing of selected objects in working memory. The SPCN typically follows the N2pc at around 350-400 ms after search display onset and has a similar topography to the N2pc. This component is assumed to be linked to the attentional activation

of visual working memory representations of potential target objects, analogous to the contralateral delay activity (CDA) that is observed during the delay period of visual working memory tasks (e.g., Vogel & Machizawa, 2004).

The roles of feature-based and object-based target templates during visual search can be investigated in conjunction search tasks by measuring ERPs to targets and to partially target-matching but task-irrelevant objects. If the guidance of visual selection was entirely feature-based, all objects with target-matching features should trigger N2pc components. If it was object-based, N2pcs should only be elicited by fully template-matching target objects, but not by nontargets with one target feature. Recent ERP studies from our lab that measured attentional biases during search for feature conjunctions (colour/shape: Eimer & Grubert, 2014; Jenkins, Grubert, & Eimer, 2017; colour/colour: Berggren & Eimer, 2016a; colour/alphanumeric category: Nako, Grubert, & Eimer, 2016) obtained converging results. In these experiments, targets (e.g., blue circles) and partially-matching nontargets (e.g., red circles and blue squares) both elicited reliable N2pc components, and the target N2pc was initially identical to the sum of the two N2pc components to each possible partially-matching nontarget. This suggests that attention was guided independently by search templates for each target-defining feature, and that these templates operated in parallel, producing additive attentional biases of visual processing. However, from about 250 ms after search display onset, the target N2pc became larger than the sum of the N2pcs to partially-matching nontargets, indicating that attentional processing was now sensitive to the joint presence of multiple target-defining features within the same object. These findings demonstrate that the selectivity of template-guided attentional modulations during conjunction search evolves across time, but do not yet provide unequivocal evidence for an involvement of object-based target templates. The emergence of superadditive target N2pcs in these earlier studies could mark the point in time when attentional selectivity starts to be guided by integrated object templates, but it could also be the result of facilitatory interactions between feature-specific attentional modulations that happen to be triggered at the same spatial coordinates within visual cortical maps (e.g., Desimone & Duncan, 1995). In this case, a superadditive target N2pc would simply reflect an enhanced attentional bias for two target-defining features at the same location (e.g.,

“blue” and “circle”), but not an object-selective attentional response (“blue circle”) that is triggered by a match with an object-based target template.

These two alternative possibilities could not be distinguished in these earlier ERP studies of conjunction search, because only targets had both template-matching features. To dissociate the effects of feature-based and object-based templates, we designed a search task where targets were defined by colour/shape conjunctions, and some search displays included objects with a target colour and a target shape that had to be recognised as nontargets. In Experiment 1, there were two possible target objects (e.g., blue circles and green squares). Participants’ task was to report the presence or absence of one of these targets on each trial. Targets were always accompanied by three non-matching distractor objects. On some target-absent trials, search displays contained a partially target-matching object with one of the two target colours or shapes among non-matching distractors. Critically, on other target-absent trials, one of two possible “incorrect conjunction” nontarget objects was presented. These objects had both a target colour and a target shape, but the combination of these two features did not match either target object (e.g., blue squares or green circles; see Figure 1, top panel). Because targets and incorrect conjunction objects both had a target-matching colour and a target-matching shape, these two types of objects could not be distinguished by independently operating feature-based attentional guidance processes. Thus, any attentional bias that is selectively triggered by targets but not by incorrect conjunction objects will reflect an involvement of genuinely object-based target templates.

EEG was recorded during task performance, and lateralised ERP components (N2pc and SPCN) were measured for targets, partially-matching nontargets, and, critically, incorrect conjunction objects, to track the time course of template-guided attentional selection and the access to visual working memory for these different types of objects. Based on our earlier findings from conjunction search tasks (see above), it was predicted that all objects with target-matching features would initially trigger N2pc components, reflecting an early feature-based phase of attentional selectivity. The critical question was whether and when attentional modulations of visual processing would become sensitive to the difference between genuine target objects and incorrect conjunction nontargets. Because both types of objects had a

target-defining colour and shape, the N2pc to these objects should initially be identical. If the guidance of attention in this two-target conjunction search task was exclusively feature-based, lateralised ERP waveforms to targets and incorrect conjunction nontargets should remain identical throughout the N2pc time interval, and possibly also during the subsequent SPCN period. Such an observation would demonstrate that visual selection and access to visual working memory are controlled independently by target colours and shapes, and are not sensitive to the presence versus absence of a target-matching feature combination. In contrast, if object-based templates were involved in the guidance of attention, target N2pc should eventually become larger than the N2pc to incorrect conjunction nontargets. In this case, the moment where this divergence occurs would mark the point in time where attentional control processes start to become genuinely object-based. It is also possible that object-based target templates have no effect on attentional modulations of visual processing, as reflected by N2pc components, but operate at a later stage where objects are encoded into working memory. In this case, targets and incorrect conjunction nontargets should trigger identical N2pc, but only target objects should elicit SPCN components.

To be able to create nontarget objects with incorrectly combined task-relevant colours and shapes, we needed to use a task where participants had to search for one of two different colour/shape conjunction targets. The control of this type of two-object search raises important additional questions on its own. For example, there is considerable debate about whether observers can maintain more than a single target template at any given time (e.g., Olivers, Peters, Houtkamp, & Roelfsema, 2011; Beck, Hollingworth, & Luck, 2012; see below for more detailed discussion). If participants were able to activate only a single template for one of the two target objects on each trial, this should affect target N2pc amplitudes, as target objects would match this template on only half of all trials. To assess this possibility, Experiment 1 also included a single-target control task where observers searched for only one colour/shape conjunction target. Because participants had to compare each selected object to only one memorised target item in this task, but to two different items in the double target task, responses should be slower in the latter task (Cunningham & Wolfe, 2014). Furthermore, this

difference in the demands on memory matching processes might also be reflected by differences in the amplitude and/or duration of SPCN components between the two tasks.

EXPERIMENT 1

Method

Participants

Twelve participants were recruited for Experiment 1 (M age = 30 years, SD = 6; 6 male; 2 left-handed). All reported normal or corrected-to-normal vision. Desired sample size was a priori defined at 12 participants, consistent with many previous N2pc and SPCN studies in our lab and elsewhere (e.g., Berggren & Eimer, 2016a; Luck & Hillyard, 1994). To assess the sufficiency of this sample size, we assessed the effect size from two previous studies from our group that were conceptually closest to the current study, which examined differences in N2pc response to targets defined by a configuration of two features (i.e., versus objects containing relevant features but task-irrelevant due to their specific configuration; Berggren & Eimer, 2016b, Experiment 3; McCants, Berggren, & Eimer, in press, Experiments 1 and 2). Across these experiments, effect sizes (based on Cohen's d ; mean difference divided by pooled standard deviation method) for differences in N2pc mean amplitudes for a target alone relative to a target in competition with a nontarget in an incorrect configuration were used as the basis for power analyses, where power of 0.8 and an alpha level of 0.05 was assumed. The relevant effect sizes ranged from 1.25 to 1.61 in these experiments. These findings suggest that a minimum sample size of 6-8 participants was required to demonstrate effects of feature configurations on N2pc amplitudes.

Stimuli and Procedure

The experimental task was created and executed using E-Prime 2.0 software (Psychology Software Tools, Inc.). All stimuli were presented on a 24-inch BenQ monitor (60 Hz; 1920 x 1080 screen resolution) at a viewing distance of approximately 90 cm. The experiment was run on a SilverStone PC, and manual responses were registered through a standard PC keyboard. All stimuli were shown on a black background, with a grey fixation dot ($0.2^\circ \times 0.2^\circ$ of visual angle) present throughout blocks. Stimuli shown within search displays were coloured shapes. The colours used in the experiment were red (CIE colour coordinates: .605/.322), orange (.543/.409), green (.296/.604), blue (.169/.152), magenta (.270/.134), or grey (.305/.325). All colours were equiluminant (14 cd/m^2). The possible shapes used were square, circle, hexagon, gate, heart, and triangle (size: $1.40^\circ \times 1.40^\circ$). Each search display contained four coloured shape objects, one in each quadrant, at an eccentricity of 2.42° from fixation (measured relative to the centre of each shape).

Participants completed two tasks. In the *double target* search task, they had to find one of two possible conjunctively-defined colour/shape targets (e.g., blue circle or green square). They responded by pressing the '1' key on the numeric keypad if one of these two targets was present or the '2' key if no target was detected. Search displays appeared for 100 ms followed by a 1900 ms blank period prior to the next search display, and responses were recorded during the full 2000 ms following search display onset. Each block began with a 1000 ms blank screen period prior to the first search display. Each search display contained three objects without target-matching features and one object that matched one or both target features (see Figure 1, upper panel). On target-present trials, this object was one of the two targets. On target-absent trials, this object either matched a target colour (e.g., blue gate), a target shape (e.g., magenta circle), or both a target colour and shape but in an incorrect conjunction (e.g., blue square). Participants were instructed to respond 'present' only to exact target matches, and to respond 'absent' to partially-matching objects and to incorrect conjunctions. Following practice, the task contained nine experimental blocks, each consisting of 80 trials. Each block contained 32 target-present and 48 target-absent trials (16 trials each with a partial colour-match, partial shape-match, and an incorrect conjunction object). Target-absent trials were more likely than target-present trials to obtain a sufficient number of trials for ERP analysis for each of the three

different types of template-matching nontarget objects. The object with target-matching feature(s) appeared equally often at each of the four possible positions. On target-absent trials, the four possible partial colour and partial shape matches and the two possible incorrect conjunction objects were all equiprobable. The *single target* search task was very similar, except that only a single colour-shape conjunction served as target, and there were therefore no incorrect conjunction objects. Participants completed nine experimental blocks consisting of 64 trials each. Each block contained 32 target-present and 32 target-absent trials (16 trials each with a partial colour-match and a partial shape-match). To avoid low-level pop-out of objects with a unique shape attribute, each search display always contained two angular (square, hexagon, or triangle) and two curved objects (circle, gate, heart) in both tasks.

Participants completed both tasks within a single session, with task order counterbalanced across participants. One training block preceded the first experimental block for each task. The choice of targets was randomised across participants, except that two colours (orange and grey) and two shapes (triangle and heart) were never used as target features for any participant. The assignment of the remaining four colours and shapes as target features in the two tasks was randomised across participants, with the restriction that there was no feature overlap between any of the targets employed for each participant in the single and double target search tasks.

EEG Recording and Data Analysis

EEG was DC-recorded at 27 scalp electrodes, mounted on an elastic cap at sites Fpz, F7, F8, F3, F4, Fz, FC5, FC6, T7, T8, C3, C4, Cz, CP5, CP6, P9, P10, P7, P8, P3, P4, Pz, PO7, PO8, PO9, PO10, and Oz. A 500-Hz sampling rate was used with a 40 Hz low-pass filter applied. Channels were referenced online to a left-earlobe electrode, and re-referenced offline to an average of both earlobes. No other filters were applied. Trials with eye blinks (exceeding ± 60 μV at Fpz), horizontal eye movements (exceeding ± 30 μV in the HEOG channels), and muscle movement artifacts (exceeding ± 80 μV at all other channels) were removed as artefacts, in addition to trials with incorrect trial responses. The average number of trials removed from the EEG data

set was 22% (SD = 11). All remaining trials were segmented into epochs from 100 ms before to 600 ms after search display onset, relative to a 100 ms pre-stimulus baseline. Averaged ERP waveforms were computed separately for the two search tasks (single/double target), for displays containing a target, a partial colour-match, a partial shape-match, or an incorrect conjunction object (for the double target task only) on the left or right of fixation. N2pc amplitudes were calculated based on ERP mean amplitudes obtained at lateral posterior electrode sites PO7 and PO8 between 200 and 300 ms post-display onset. SPCN amplitudes were quantified based on ERP mean amplitudes at the same electrode sites between 400-600 ms post-display onset.

Results

Behavioural performance

Figure 2 (left panel) shows mean reaction times (RTs) and error rates for the different display types in both tasks of Experiment 1. RTs on trials with correct responses were initially analysed separately for the two search tasks. A one-way repeated-measures Analysis of Variance (ANOVA) for the double target task showed a main effect of Display Type (target-present, colour-match, shape-match, incorrect conjunction; $F(3,33) = 156.73$, $p < .001$, $\eta_p^2 = .93$). Pairwise comparisons showed that RTs differed significantly between each of the four different display types (t 's > 4.09 , p 's $< .005$, d 's > 1.18). RTs were fastest for shape-match displays ($M = 584$ ms), slower for colour-match ($M = 616$ ms) and target-present displays ($M = 646$ ms), and were strongly delayed with incorrect conjunction displays ($M = 757$ ms). In the single target task, the factor Display Type (target-present, colour-match, shape-match) also showed a significant main effect ($F(2,22) = 6.24$, $p = .007$, $\eta_p^2 = .36$). Pairwise comparisons revealed that RTs were significantly slower for colour-match displays ($M = 535$ ms) as compared to both target-present ($M = 506$ ms; $t(11) = 2.71$, $p = .02$, $d = .78$) and shape-match displays ($M = 505$ ms; $t(11) = 4.20$, $p = .001$, $d = 1.21$). RTs did not differ significantly between target-present and shape-match displays ($t < 1$). When collapsed across all display types, RTs were generally

slower in the double target task relative to the single target task ($M = 651$ ms vs. 515 ms; $t(11) = 12.27$, $p < .001$, $d = 3.54$).

Matching analyses were conducted on error rates. A significant effect of Display Type on error rates was present in the double target task ($F(3,33) = 14.04$, $p < .001$, $\eta_p^2 = .56$). Error rates were higher for target-present displays and for displays containing an incorrect conjunction object ($M = 6\%$ and 12% respectively) than for colour-match and shape-match displays ($M = 1\%$ and 0% respectively; t 's > 3.78 , p 's $< .005$, d 's > 1.09). The difference between error rates for target-present versus incorrect conjunction displays only approached significance ($t(11) = 2.04$, $p = .066$, $d = .59$), and there was no error rate difference between displays with colour versus shape-matches ($t < 1$). In the single target task, there was also a significant effect of Display Type ($F(2,22) = 6.61$, $p = .006$, $\eta_p^2 = .38$), as participants missed targets more frequently on target-present trials ($M = 4\%$) than incorrectly reporting the presence of a target in displays with a colour-match or shape-match ($M = 2\%$ and 1% ; t 's > 2.22 , p 's $< .05$, d 's $> .64$). The difference in error rates between colour-match and shape-match displays missed significance ($M = 2\%$ vs. 1% ; $t(11) = 1.92$, $p = .08$, $d = .55$). Finally, when collapsing across display types, error rates were significantly higher in the double target versus single target task ($M = 5\%$ vs. 2% ; $t(11) = 4.06$, $p = .002$, $d = 1.17$).

ERP results

N2pc components. Figures 3 and 4 show grand-averaged ERPs triggered at electrodes PO7/8 contralateral and ipsilateral to the side of the object with target-matching features. ERPs are shown separately for the different display types in the single and double target tasks, together with the corresponding contralateral-ipsilateral difference waveforms. ANOVAs of ERP mean amplitudes obtained 200-300 ms post-stimulus were conducted separately for both tasks, with the factors Display Type and Laterality (Ipsilateral, Contralateral). In the double target task, a main effect of Laterality ($F(1,11) = 34.64$, $p < .001$, $\eta_p^2 = .76$) was accompanied by a significant interaction with Display Type ($F(3,33) = 16.49$, $p < .001$, $\eta_p^2 = .60$). Significant N2pc components were found for all four display types (t 's > 3.96 , p 's $< .005$, d 's > 1.16). As can be

seen in Figure 3, the N2pc to shape-matches (M diff = $-0.47 \mu\text{V}$) was strongly attenuated relative to the N2pc components elicited by the other types of displays (t 's > 4.17 , p 's $< .005$, d 's > 1.20). Importantly, there was no difference in the size of N2pc components triggered by targets and by incorrect conjunction objects (M diff = $-2.15 \mu\text{V}$ vs. $-2.20 \mu\text{V}$; $t < 1$). Relative to these two N2pcs, the N2pc component elicited by colour-matching nontarget objects was numerically smaller (M diff = $-1.84 \mu\text{V}$), but these differences did not reach statistical significance (t 's < 1.89 , p 's $> .08$, d 's $< .55$).

In the single target task (Figure 4), a significant main effect of Laterality ($F(1,11) = 19.31$, $p = .001$, $\eta_p^2 = .64$), indicating the presence of reliable N2pc components, was accompanied by an interaction between Laterality and Display Type ($F(2,22) = 8.91$, $p = .001$, $\eta_p^2 = .45$). Pairwise comparisons between ipsilateral and contralateral ERP amplitudes showed that significant N2pc components were elicited in response to all three types of displays (t 's > 2.55 , p 's $< .03$, d 's $> .73$), but that target objects elicited a significantly larger component (M diff = $-2.17 \mu\text{V}$) relative to colour-matches (M diff = $-1.43 \mu\text{V}$; $t(11) = 3.34$, $p = .007$, $d = .96$) and shape-matches (M diff = $-.58 \mu\text{V}$; $t(11) = 4.39$, $p = .001$, $d = 1.27$). There was no significant difference between N2pc mean amplitudes elicited by colour versus shape-matches ($t(11) = 1.38$, $p = .20$, $d = .40$).

SPCN components. Mean amplitudes obtained 400-600 ms post-stimulus onset were entered into separate ANOVAs for each task. In the double target task, there was a main effect of Laterality ($F(1,11) = 19.84$, $p = .001$, $\eta_p^2 = .64$) that interacted with Display Type ($F(3,33) = 13.24$, $p < .001$, $\eta_p^2 = .55$). As can be seen in Figure 3, targets as well as incorrect conjunction objects both elicited clear SPCN components (M diff = $-1.21 \mu\text{V}$ vs. $-1.43 \mu\text{V}$; t 's > 4.78 , p 's $\leq .001$, d 's > 1.37), while colour-matching and shape-matching nontargets did not (M diff = $-.15 \mu\text{V}$ vs. $-.13 \mu\text{V}$; t 's < 1). Importantly, there was no difference in SPCN amplitudes between targets and incorrect conjunction objects ($t < 1$). In the single target task (Figure 4), the main effect of Laterality failed to reach significance ($F(1,11) = 2.58$, $p = .14$, $\eta_p^2 = .19$), but this was qualified by a significant interaction with Display Type ($F(2,22) = 6.92$, $p = .005$, $\eta_p^2 = .39$). Comparing ipsilateral versus contralateral ERP amplitudes, a reliable SPCN was only elicited by

target objects (M diff = $-0.68 \mu\text{V}$; $t(11) = 3.02$, $p = .012$, $d = .88$), but not by colour- or shape-matching nontargets (M diff = $.24 \mu\text{V}$ vs. $-.14 \mu\text{V}$; t 's < 1.94 , p 's $> .05$, d 's $< .55$).

Comparisons between tasks. N2pc components triggered by target objects did not differ in size between the single target and double target tasks (M diff = $-2.17 \mu\text{V}$ vs. $-2.15 \mu\text{V}$; $t < 1$). In contrast, SPCN amplitudes triggered by target objects were significantly larger in the double target as compared to the single target task (M diff = $-1.21 \mu\text{V}$ vs. $-.68 \mu\text{V}$; $t(11) = 4.49$, $p = .001$, $d = 1.30$).

Discussion of Experiment 1

The results of the double target task provided no evidence for any involvement of object-based target templates in the guidance of conjunction search. Reliable N2pc components were elicited by colour-matching and shape-matching nontargets, demonstrating that attentional selectivity operated in a feature-based fashion. The critical finding was that targets and incorrect conjunction objects triggered identical N2pcs in the double target task, demonstrating that attentional guidance operated in an exclusively feature-based fashion. The pattern of behavioural and N2pc results obtained in this task suggest that these guidance processes were dominated by colour signals. N2pc amplitudes to colour-matching nontargets in the double target task were reliably larger than to shape-matching nontargets. A similar difference was also present in the single target task, although it was not statistically significant. These observations are in line with previous N2pc studies of conjunction search where N2pcs were consistently larger when a nontarget object matched the target colour than when it had a different target-defining attribute (e.g., Eimer & Grubert, 2014; Nako et al., 2016). This suggests that colour signals are generally more effective in guiding attention than target-matching features from other dimensions. Colour-based attentional guidance may dominate guidance by other features particularly when the memory load associated with storing target templates is

high, as is the case when two conjunctively defined objects both serve as targets. However, the fact that N2pcs were not only triggered by colour-matching nontargets but were also reliably present for shape-matching nontargets in both tasks of Experiment 1 shows that even in the double target task, attentional guidance was not exclusively colour-based.

While significant N2pcs were triggered by all objects with target-matching features in the double target task, the subsequent SPCN component was elicited only by objects with both target-defining features in the same object (i.e., targets and incorrect conjunction nontargets). The absence of SPCN components for colour-matching or shape-matching objects indicates that objects with only one target feature were not encoded into working memory. It is notable that this was also the case for colour-matching nontargets, which elicited large N2pc components, because it suggests that access to visual working memory was regulated by interactions between colour-based and shape-based guidance processes. The fact that incorrect conjunction objects elicited clear SPCN components shows that these objects were maintained as candidate target objects in working memory. This conclusion is also supported by the fact that target-absent RTs were delayed by more than 100 ms when an incorrect conjunction object was present. Most importantly, there were no reliable SPCN amplitude differences between targets and incorrect conjunction nontargets, which strongly suggests that access to working memory was regulated in an entirely feature-based fashion, and that object-based target templates were not involved. Overall, the ERP results from the double target task of Experiment 1 demonstrate that for at least 600 ms after search display onset, attentional modulations of perceptual and memory-related processing were not sensitive to the difference between genuine targets and incorrect conjunction objects.

If only a single target template can be active at any moment (e.g., Olivers et al., 2011), participants may have been able to maintain only one colour and one shape template on a given trial. In this case, the N2pc to target objects in the double target task should have been smaller than the target N2pcs in the single target control task of Experiment 1, where all target objects should have matched the currently active feature templates. In fact, target N2pc amplitudes in these two tasks were virtually identical. This suggests that it is possible to simultaneously maintain feature-specific templates for two different target colours and two

different target shapes, and that the attentional biases triggered by these templates are similar to the biases elicited when only one colour and one shape template are active (see also Moore & Weissman, 2010; Beck et al., 2012; Irons, Folk, & Remington, 2012; Grubert & Eimer, 2016 for further behavioural and ERP evidence for attentional guidance by multiple target features). However, RTs to targets were delayed in the double target relative to the single target task. As target N2pc components did not differ between these two tasks, this delay is likely to be generated not during the initial allocation of attention to particular objects, but at a subsequent stage where selected objects are encoded into working memory and are compared to the memory set of current target objects. The duration of this memory matching process is assumed to increase as the number of possible target objects increases (Cunningham & Wolfe, 2014). It is notable that SPCN amplitudes to target objects was nearly twice as large in the double target task where a candidate target object had to be matched with two stored target object representations, relative to the single target task where an attended object had to be compared with only one target representation. This suggests that the SPCN does not just reflect the encoding of an attended object into working memory, but also the matching process between this object and the stored set of target objects, with larger amplitudes for two versus one memory matching processes. In other words, the SPCN might be an electrophysiological correlate of the memory search stage postulated by Cunningham and Wolfe (2014).

While the results of Experiment 1 suggest that attentional templates operate in an entirely feature-based fashion during conjunction search, such a general conclusion should not be based on the outcome of a single study. It remains possible that object-based target templates are available for the guidance of search, but that these templates are only employed when this is required by the demands of a specific selection task. In Experiment 1, search displays always included only a single object with target-matching features. Under these conditions, the candidate target object in each display could always be selected on the basis of feature-based templates, and then be encoded into working memory for identification, without any need for additional object-based attentional guidance. Experiment 2 was conducted to test whether object-based attentional templates are activated to guide attention in a task context where feature templates are no longer sufficient to localise target objects.

EXPERIMENT 2

In Experiment 2, all search display objects contained at least one target-matching feature. Participants again searched for one of two target objects defined by colour/shape conjunctions. These targets now appeared either together with three partially-matching nontargets, or with two partially-matching nontargets and one incorrect conjunction object. Target-absent displays contained one or two incorrect conjunction objects among partially-matching nontargets (see Figure 1, bottom panel). In all displays, two stimuli were presented on the left and right side and two on the vertical meridian (above and below fixation). This was done to measure N2pc and SPCN components to a target or an incorrect conjunction nontarget object in displays where both of these objects were present, and one of them appeared on the vertical midline. Because no lateralised N2pc or SPCN is elicited by vertical objects, these components exclusively reflect the attentional processing of candidate target objects on the left or right side (e.g., Woodman, & Luck, 2003; Hickey, McDonald, & Theeuwes, 2006; Eimer, Kiss, & Nicholas, 2011; Eimer & Grubert, 2014). N2pc and SPCN components were measured for targets and incorrect conjunction objects, separately for no-competition trials where they were presented together with three partially-matching objects and for competition trials where they appeared in the same display.

Because all search display objects now contained target-matching features, they should all trigger feature-based attentional biases, thus increasing the competition for attentional selection in all displays. The presence of one or two target-colour nontarget objects in each display should also make it less likely that feature-based guidance was dominated by colour (as was the case in Experiment 1), as colour signals were now less useful to indicate the location of candidate target objects. Most importantly, because a target and an incorrect conjunction object could appear in the same display in Experiment 2, participants could no longer rely exclusively on feature-based guidance processes to select one candidate target object. The critical question was whether object-based target templates would now be involved in the guidance of attention. One possibility is that object-based templates remain unavailable to

attentional control processes even when these templates would be helpful to localise target objects. In this case, attentional selectivity should remain entirely feature-based, reflected by identical N2pc and SPCN components to targets and incorrect conjunction objects, as in Experiment 1. Alternatively, if object-based templates were activated, lateralised ERPs to targets and incorrect conjunction objects should diverge at some point in time, with larger N2pc and/or SPCN components for target objects. To preview the main results, this is exactly what was observed in Experiment 2. However, additional analyses based on RT median splits revealed that even though participants had to search for two possible target objects, only a single object-based target template was activated on any given trial.

Method

Participants

Fifteen participants took part in Experiment 2. Three participants were excluded from analysis due to excessive HEOG artefact rejection (> 50 % of data). The remaining 12 participants (M age = 31 years, SD = 6; 5 male; 1 left-handed) all reported normal or corrected-to-normal vision. To confirm that Experiment 1 had been appropriately powered, we assessed all effect sizes for significant N2pc mean amplitude differences in both the previous single and double target task. Cohen's d effect sizes ranged from .98 to 1.47, suggesting a minimum sample size of between 6-11 participants, and demonstrating that Experiment 1 was not underpowered. We therefore used the same sample size (N = 12) in Experiment 2, also to ensure comparable signal-to-noise ratios of the grand-averaged ERP waveforms between these two experiments.

Stimuli and Procedure

Stimuli and procedure were similar to Experiment 1, with the following exceptions. The four object locations within each search display were now directly above, below, left, and right of fixation (centre of each shape's distance to fixation: 1.66° eccentricity; see Figure 1, lower

panel). Participants now completed only a double target search task. Both possible targets were again defined by a combination of colour and shape. All display objects contained either one or two target-defining features.

Each block included 72 trials. On 48 trials, a target was present and was equally likely to appear on the horizontal midline (to the left or right of fixation) or on the vertical midline (above or below fixation). The target appeared together with an incorrect conjunction object and one colour-matching and one shape-matching nontarget on 32 trials (8 trials for each of the four possible combinations of target location: horizontal/vertical, and incorrect conjunction object location: horizontal/vertical). On the remaining 16 target-present trials, the target was accompanied by three partially-matching nontargets (one colour-match and two shape-matches, or vice versa), and appeared either on the horizontal or vertical midline. There were 24 target-absent trials in each block. On half of these trials, an incorrect conjunction object appeared either on the horizontal or vertical midline, together with three partially-matching objects (one colour-match and two shape-matches, or vice versa). In the remaining 12 target-absent trials, displays contained the two different incorrect conjunction objects, one on the left and one on the right side. Trials where search displays contained two objects with both target-matching features were labelled competition trials, and trials where a target or an incorrect conjunction object appeared with three partially-matching nontargets were labelled no-competition trials. An additional constraint was that no display could contain three objects with the same colour or shape, to avoid any low-level pop-out for the remaining object. As in Experiment 1, participants responded to the presence or absence of either target object, and were instructed to ignore partial matches and incorrect conjunction objects. Following practice, participants completed 18 experimental blocks.

EEG Recording and Data Analysis

EEG recording and data analysis matched Experiment 1. The average number of trials removed prior to EEG analyses was 22% (SD = 14). Average ERP waveforms were computed for trials where a target or an incorrect conjunction object appeared on the left or right side, separately for displays where this object appeared among three partially-matching nontargets

(no-competition trials) and for displays that also included an incorrect conjunction object or a target on the vertical midline (competition trials). No ERPs were computed for displays containing a vertical target or incorrect conjunction object among three partially-matching nontargets, and displays with two lateral incorrect conjunction objects, because these displays would not trigger lateralised attentional responses. N2pc and SPCN amplitudes were again calculated based on ERP mean amplitudes obtained at PO7/PO8 from 200-300 ms and 400-600 ms after display onset, respectively. Three additional sets of analyses were conducted to assess the point in time when lateralised ERP responses to targets started to differ from responses to incorrect conjunction objects. First, N2pc mean amplitudes were analysed for three successive post-stimulus time windows (200-240 ms, 240-280 ms, 280-320 ms). Next, N2pc difference waves to targets and incorrect conjunction objects were compared with unidirectional one-sample *t*-tests conducted for each post-stimulus sampling point (see Jenkins et al., 2017, for analogous methods). Finally, N2pc difference waves for incorrect conjunction objects were subtracted from N2pc difference waves for targets, separately for competition and no-competition trials. The onset of differences in the resulting double subtraction waveforms was compared using a jackknife-based analysis method (see Miller, Patterson & Ulrich, 1998, for details) with an absolute onset criterion of $-0.4 \mu\text{V}$.

Results

Behavioural performance

Figure 2 (right panel) shows reaction times (RTs) and error rates for the different display types in Experiment 2. Participants responded slower to targets in displays where they were accompanied by an incorrect conjunction object than in displays where no such object was present ($M = 761$ ms vs. 832 ms). Target-absent responses were faster for displays with one as compared to two incorrect conjunction objects ($M = 945$ ms vs. 1061 ms). Correct RT data were entered into a 2x2 repeated-measures ANOVA with the factors Target Presence (present, absent) and Competition (target or incorrect conjunction object alone, target and incorrect

conjunction or two incorrect conjunction items in the same display). There were main effects of Target Presence ($F(1,11) = 158.21, p < .001, \eta_p^2 = .94$), reflecting faster RTs on target-present trials, and Competition ($F(1,11) = 34.44, p < .001, \eta_p^2 = .76$), demonstrating slower RTs to displays including two objects with both target-matching features. The interaction between these two factors did not reach significance ($F(1,11) = 3.59, p = .085, \eta_p^2 = .25$).

A matching analysis of error rates revealed significant main effects of Target Presence ($F(1,11) = 22.33, p = .001, \eta_p^2 = .67$), Competition ($F(1,11) = 22.10, p = .001, \eta_p^2 = .67$), and an interaction between both factors ($F(1,11) = 15.63, p = .002, \eta_p^2 = .59$). Participants missed targets more often when they were accompanied by an incorrect conjunction object than when no such object was present ($M = 14\% \text{ vs. } 8\%; t(11) = 5.23, p < .001, d = 1.51$). A smaller effect of competition on error rates was also found on target-absent trials, with more False Alarms for displays with two versus one incorrect conjunction objects ($M = 6\% \text{ vs. } 4\%; t(11) = 2.38, p = .04, d = .69$).

ERP results

N2pc components. Figure 5 shows grand-averaged ERPs triggered at electrodes PO7/8 contralateral and ipsilateral to the side of a target object or an incorrect conjunction object on trials where these objects appeared on the left or right side. ERPs are shown separately for no-competition trials where displays did not contain another object with both target-matching features, and for competition trials where a target and an incorrect conjunction object appeared in the same display (one horizontal, one vertical). The corresponding contralateral-ipsilateral difference waveforms are also shown. In both competition and no-competition trials, targets and incorrect conjunction objects initially triggered N2pc components of similar size. However, and in contrast to Experiment 1, the target N2pc became larger than the N2pc to incorrect conjunction objects from about 280 ms after display onset.

In an initial analysis, ERP mean amplitudes obtained at PO7/8 during the 200-300 ms post-stimulus interval were entered into a 2x2x2 ANOVA with the factors Object Type (lateral target, lateral incorrect conjunction object), Competition (no-competition displays: target or

incorrect conjunction object alone; competition displays: target and incorrect conjunction object both present), and Laterality. A significant main effect of Laterality ($F(1,11) = 27.92, p < .001, \eta_p^2 = .72$) confirmed the reliable presence of N2pc components. There was no significant Competition x Laterality interaction ($F(1,11) = 1.99, p = .19, \eta_p^2 = .15$), indicating that the presence or absence of a competing object with two target-matching features did not affect the size of these components. More importantly, there was also no reliable interaction between Object Type and Laterality ($F(1,11) = 2.86, p = .12, \eta_p^2 = .21$), suggesting that across the whole 200 – 300 post-stimulus time window, the N2pc to targets was not reliably larger than the N2pc to incorrect conjunction objects. There was also no three-way interaction between all factors ($F(1,11) = 1.96, p = .19, \eta_p^2 = .15$). Pairwise comparisons between ipsilateral and contralateral ERP amplitudes showed that significant N2pc components were elicited in response to all four types of displays shown in Figure 5 (target alone: $-.98 \mu\text{V}$; incorrect conjunction item alone: $-.94 \mu\text{V}$; lateral target and vertical incorrect conjunction object: $-.93 \mu\text{V}$; lateral incorrect conjunction object and vertical target: $-.58 \mu\text{V}$; all t 's > 3.32 , all p 's $< .01$, d 's $> .96$).

The divergence between N2pc components to targets and incorrect conjunction objects that emerged around 280 ms after display onset both for competition and no-competition displays (Figure 5) could suggest that object-based attentional guidance emerges after an early strictly feature-based phase of attentional selectivity. To assess this statistically, N2pc mean amplitudes (collapsed across no-competition and competition trials) were analysed separately across three successive 40 ms time window (200-240, 240-280, and 280-320 ms post-stimulus). An initial 3x2x2 ANOVA including the factor Time Window showed a significant three-way interaction ($F(2,22) = 4.68, p = .02, \eta_p^2 = .30$). Follow-on analyses conducted separately for each time interval showed that there were no Object Type x Laterality interactions in the first two intervals (F 's $< 1.99, p$'s $> .19, \eta_p^2$'s $< .15$), indicating that N2pcs to targets and incorrect conjunction objects did not differ during these periods. In contrast, this interaction was significant for the third interval ($F(1,11) = 5.46, p = .039, \eta_p^2 = .33$), where a reliable N2pc was present only for targets ($M \text{ diff} = -.98 \mu\text{V}; t(11) = 3.22, p = .008, d = .93$) but not for incorrect conjunction objects ($M \text{ diff} = -.38 \mu\text{V}; t(11) = 1.80, p = .10, d = .52$).

These N2pc differences between targets and incorrect conjunction objects suggest that object-based target templates started to modulate attentional biases during the late phase of the N2pc. However, they do not reveal whether templates for both target objects were simultaneously activated on each trial. If only a single object-based template can be maintained at any given time, this template would match the target on half of all target-present trials. On the other half of trials, a mismatch between the target and the currently active object template would prevent any object-based guidance, resulting in attentional biases that are entirely feature-based. To assess this possibility, we performed an RT median split for trials with lateral target objects (collapsed across competition and no-competition trials), individually for each participant, and computed ERPs separately for trials with fast and slow responses. If only one object template was active at any time, trials with fast target-present RTs should primarily be trials where the target matched this template, while slow RTs should be the result of a mismatch between the current target and the object template. In this case, the target N2pc on fast RT trials should be larger than the target N2pcs on trials with slow RTs, which should be similar to the N2pc to incorrect conjunction objects. This was exactly what was observed. Figure 6 shows N2pc difference waveforms for target-present trials with fast and slow RTs and for incorrect conjunction objects. An initial analysis of ERP mean amplitudes in the N2pc time window (200-300 ms post-stimulus) with the factors Laterality and Trial Type (target-present, fast RTs; target-present, slow RTs; incorrect conjunction object) showed an interaction between both factors ($F(2,22) = 7.16, p = .004, \eta_p^2 = .39$). Follow-up paired t-tests confirmed that target N2pcs on trials with fast RTs were larger than both target N2pcs on slow RT trials (M diff = $-1.26 \mu\text{V}$ vs. $-.67 \mu\text{V}$; $t(11) = 3.07, p = .011, d = .89$) and N2pcs to incorrect conjunction objects (M diff = $-0.80 \mu\text{V}$; $t(11) = 2.67, p = .022, d = .77$). There was no amplitude difference between the target N2pcs on slow RT trials and the N2pc to incorrect conjunction objects ($t(11) = 1.08, p = .31, d = .31$).

If the N2pc on target-present trials with fast RTs reflects attentional biases that are affected by an object-based template, the moment where such a template becomes active can be tracked by analysing when reliable N2pc differences between these trials and trials with incorrect conjunction objects start to emerge. We tested this in three separate analyses that

produced converging results. First, we compared the amplitudes of these two N2pc components within successive 40 ms time windows, as described above. There was no Object Type x Laterality interaction within the first time window ($F(1,11) = 1.92, p = .19, \eta_p^2 = .15$), but this interaction was reliable within the two later time windows (240-280 ms and 280-320 ms post-stimulus; F 's $> 5.11, p$'s $< .05, \eta_p^2$'s $> .32$), reflecting larger N2pc amplitudes to targets on trials with fast RTs than to incorrect conjunction objects that emerged at or shortly after 240 ms after search display onset. Next, a series of unidirectional one-sample t -tests were conducted for each post-stimulus sampling point to compare N2pc difference waveforms for targets on fast RT trials and incorrect conjunction objects. The onset of N2pc amplitude differences between these two N2pcs was defined as the point in time where a significant difference emerged that remained reliable for at least ten successive sampling points. This procedure yielded an onset estimate of 250 ms post-stimulus. Finally, a jackknife-based analysis with an absolute onset criterion of $-0.4 \mu\text{V}$ was conducted for double-subtraction N2pc difference waveforms (obtained by subtracting the N2pc difference waveform for incorrect conjunction objects from the N2pc difference wave for targets on trials with fast RTs). This analysis again suggested that the target N2pcs started to become larger than N2pcs to incorrect conjunction objects at 250 ms after search display onset.

SPCN components. Mean ERP amplitudes at PO7/8 obtained 400-600 ms post-stimulus were entered into an ANOVA with the factors Object Type, Competition, and Laterality. The main effect of Laterality only approached significance ($F(1,11) = 4.27, p = .063, \eta_p^2 = .28$). There was no interaction between Competition and Laterality and no three-way interaction between all factors (F 's < 1). However, and critically, the interaction between Object Type and Laterality was significant ($F(1,11) = 4.85, p = .05, \eta_p^2 = .31$). As can be seen in Figure 5, target objects elicited clear SPCN components on both competition and no-competition trials, whereas the SPCN to incorrect conjunction objects was small or absent. To assess this, pairwise comparisons between ipsilateral and contralateral ERP mean amplitudes in the SPCN time window were conducted for targets and for incorrect conjunction objects (collapsed across competition and no-competition trials). Target objects elicited reliable SPCN components (M diff = $-.63 \mu\text{V}; t(11)$

= 2.37, $p = .037$, $d = .68$), whereas incorrect conjunction objects did not (M diff = $-.22 \mu\text{V}$; $t(11) = 1.23$, $p > .20$, $d = .35$). As shown in Figure 6, SPCN components to targets appear to be present both on trials with fast and slow RTs. A direct comparison of SPCN mean amplitudes on these two types of trials revealed no significant difference (M diff = $-.58 \mu\text{V}$ vs. $-.75 \mu\text{V}$; $t < 1$), and additional analyses confirmed that target SPCNs were present on both fast and slow RT trials (t 's > 2.14 , p 's $< .03$, one-tailed, d 's $> .62$).

Discussion of Experiment 2

The ERP results of Experiment 2 provide novel evidence that object-based attentional templates are involved in the guidance of attention during conjunction search under conditions where these templates are required to distinguish between target and nontarget objects. Importantly, they also suggest that only one object-based target template can be active at any given time. In contrast to Experiment 1, where targets and incorrect conjunction objects elicited identical N2pc and SPCN components, lateralised ERP waveforms to these two types of objects diverged around 280 ms post-stimulus, with larger N2pc amplitudes to targets from this point onwards, both on no-competition and competition trials (Figure 5). Because targets and incorrect conjunction objects both had two target-defining features and this could not have been distinguished by feature-based attentional guidance processes, this divergence of N2pcs during the late phase of the N2pc time window shows that object-based control mechanisms affected the allocation of attention. The comparison of target N2pcs elicited on trials with fast and slow RTs provided additional insights into the nature of these object-based mechanisms. On trials with fast target-present responses, the target N2pc started to become larger than the N2pc to incorrect conjunction objects from 250 ms post-stimulus onwards. In contrast, target N2pc amplitudes on trials with slow responses did not differ from N2pc amplitudes to incorrect conjunction objects (Figure 6). This pattern of results suggests that participants were able to maintain only one object-based target template on any given trial. When this template matched the current target, RTs were fast, and object-based attentional control affected visual

processing rapidly, after about 250 ms. When the target did not match the current object template, target-present RTs were slow, and attentional biases remained entirely feature-based, as shown by the fact that target N2pc components were identical to the N2pc to incorrect conjunction objects. To rule out the possibility that the larger target N2pc on fast as compared to slow RT trials was due to factors unrelated to object-based templates, such as general inattention to the search display on trials with slow RTs, we ran an analogous analysis of N2pc amplitudes for incorrect conjunction objects. There were no systematic N2pc differences between fast and slow RT trials for these objects, suggesting that the differences observed for target N2pc were indeed associated with a mismatch between a particular target and the currently active object template.

In contrast to Experiment 1, where identical SPCN components were elicited by targets and incorrect conjunction objects, only targets elicited reliable SPCN components in Experiment 2, which demonstrates that access to working memory was no longer feature-based, but was controlled in an object-based fashion. However, results also suggest that there may be differences between the object-based guidance of attention and the object-based control of visual working memory. In contrast to target N2pc components, which showed clear amplitude differences between trials with fast and slow RTs, target objects elicited SPCN components of similar size regardless of whether responses were fast or slow (Figure 6). This dissociation suggests that while only a single object template can guide the initial allocation of attention, access to working memory may be regulated by multiple object templates.

The fact that N2pc amplitudes elicited before 250 ms post-stimulus did not differ between targets and incorrect conjunction objects in Experiment 2 indicates that up to this point in time, attention was guided in an entirely feature-based fashion, resulting in identical feature-based attentional biases for both types of objects. The fact that the presence of an incorrect conjunction object in the same display delayed target RTs and increased error rates is likely to reflect the early impact of competing feature-based biases in these displays.

GENERAL DISCUSSION

The goal of this study was to find out whether attentional guidance processes during conjunction search are exclusively feature-based, or whether integrated representations of target objects (object-based attentional templates) are also involved. Results demonstrate that object-based target templates can guide attention towards target objects, but only under conditions where purely feature-based templates are insufficient. They also show that the effects of feature-based control on visual processing emerge earlier than the effects of object-based guidance mechanisms.

In the double target task of Experiment 1, where search displays only contained a single object with target-matching features, targets and incorrect conjunction objects triggered identical N2pc and SPCN components. This demonstrates that both types of objects elicited equivalent attentional modulations of visual processing, and both gained access to working memory, suggesting that attentional control was entirely feature-based. In this task, object-based attentional templates were not required to single out one specific object as a candidate target, and the identification of such an object as target or nontarget could thus take place at a later stage once this object was encoded into working memory. This was no longer possible in Experiment 2 where all objects had target-matching features, and a target and an incorrect conjunction object could be present in the same display. Here, ERPs revealed that object-based target templates were now involved in the control of attention. Only targets but not incorrect conjunction objects elicited reliable SPCN components in Experiment 2, indicating that only targets were encoded into working memory. Critically, lateralised ERPs for these two types of objects already diverged earlier, during the N2pc time window. On target-present trials with fast RTs, target N2pc amplitudes became larger than N2pc to incorrect conjunction objects from 250 ms after display onset, demonstrating that object-based attentional control mechanisms started to affect visual processing from this point in time onwards. Notably, target N2pc on trials with slow RTs did not differ from N2pc to incorrect conjunction objects, which suggests that only feature-based attentional control processes were active on these trials. This observation provides new electrophysiological evidence that only a single object-based target

template can be active at any given time. Such a template will result in target N2pc modulations indicative of object-based attentional guidance for displays that include a template-matching target object, and purely feature-based N2pc effects for displays with a non-matching target.

These results suggest that while multiple feature-based templates are maintained concurrently, only a single object-based target template is available to guide attention at any given moment. This would qualify previous more general suggestions that visual search can be guided by a single target template only (e.g., Olivers et al., 2011). This may in fact only apply to genuinely object-based target templates but not to templates for target features. Such a difference between search templates for features versus objects may be linked to differences in the attentional demands of maintaining these two types of templates in working memory. In object-based templates, individual target features are bound and represented in an integrated fashion (as in the object files proposed by Kahneman et al., 1992). In contrast to the storage of individual features which can operate in parallel, the maintenance of feature bindings in working memory requires focal attention (Wheeler & Treisman, 2002), and may therefore be limited to one object at a time. In the current Experiment 2, where N2pc components revealed the involvement of object-based templates, participants may have activated templates for all four target-defining features (e.g., blue, green, circle, square), but only one binding between two of these features as an object template (e.g., blue circle) on any given trial.

The current study also revealed differences in the time course of the effects of feature-based and object-based attentional templates. Until 250 ms after search display onset, attentional guidance was found to be entirely feature-based, and object-based control started to affect visual processing only after this point in time. This is in line with observations from previous N2pc studies of conjunction search (Eimer & Grubert, 2014; Berggren & Eimer, 2016a; Nako et al., 2016; Jenkins et al., in press), which found that target N2pc amplitudes were initially equal in size to the sum of N2pcs to partially-matching nontargets, but became larger (i.e., superadditive) around 250 ms after search display onset. The design of these earlier experiments made it impossible to decide whether the emergence of a superadditive N2pc reflected object-based attentional control or local facilitatory interactions between feature-specific attentional biases. The current study resolved this uncertainty by demonstrating that

genuinely object-based attentional guidance processes can indeed be activated from about 250 ms post-stimulus. The temporal sequence of feature-based followed by object-based control found here could imply that feature-based and object-based target templates become available to the guidance of attention at different points in time during a search episode. For example, the activation of individual target feature representations in working memory might be faster than the activation of a binding between these features (see above). Alternatively, it is possible that both feature-based and object-based templates are activated concurrently, but affect visual processing at different times. During the earliest stages of processing a search display, only feature information may be available, and any attentional modulations will thus be entirely feature-based. At some later point, information about feature conjunctions becomes accessible to attentional control processes, resulting in the emergence of object-based attentional biases. In this case, the point in time where N2pc components first show evidence for object-based guidance could be interpreted as an electrophysiological marker for the start of feature integration processes during visual search (e.g., Treisman, 1988).

The current experiments also revealed new insights into the attentional control of working memory. According to the theory of visual attention (TVA) proposed by Bundesen and colleagues (Bundenen, 1990; Bundesen, Habekost, & Kyllingsbæk, 2005), the goal of attentional selectivity is to determine which of multiple competing objects are encoded into capacity-limited visual working memory. The competition between objects for access to memory is controlled by the pertinence (i.e., task-relevance) of their features which are used to compute the attentional weights for each individual object. The pattern of SPCN components observed in Experiments 1 and 2 suggests that the selection of objects for encoding into memory differed as a function of the competition between objects in the same display. In the double target task of Experiment 1, targets and incorrect conjunction objects both triggered SPCNs, whereas partially-matching nontargets did not. This suggests that in a situation where there was never any direct competition between targets and incorrect conjunction objects in the same display, access to working memory was controlled in a feature-based fashion. Objects were allowed to enter working memory when they possessed a target-matching colour and a target-matching shape, regardless of how these two features were combined. In line with the assumptions of

TVA, this implies that attentional weights were computed independently for different feature dimensions, without any interaction between dimensions. In contrast, reliable SPCN components were elicited only by targets but not by incorrect conjunction objects in Experiment 2 where these two types of objects could appear in the same display. This suggests that memory access was now controlled in an object-based fashion. In terms of TVA, this would imply the existence of interactive links between the pertinence of different features during the computation of attentional weights. It is also notable that in contrast to the N2pc, identical SPCN components were elicited on trials with fast and slow target-present responses, suggesting that object-based templates for both targets were involved in regulating access to working memory. The pattern of N2pc and SPCN results in Experiment 2 suggests that early attentional modulations of visual processing can be guided by only a single object-based target template, but that more than one object template can be involved in the control of working memory. An alternative possibility is that object-based target templates are activated serially, and that the second template becomes active at a time when it can no longer affect early attentional biases as reflected by the N2pc, but soon enough to affect working memory.

If the number of object-based attentional templates was limited to one (for the control of attentional guidance) and possibly two (for the control of access to working memory), how does this relate to the fact that observers are able to search for any of up to 100 possible real-world target objects (Wolfe, 2012; Wolfe & Cunningham, 2014)? This type of hybrid visual and memory search obviously depends on stored representations of multiple target objects, but the role of these representations is different from the roles of attentional target templates that were investigated in the current experiments. Attentional templates are involved in the guidance of attention towards possible target objects, and possibly also control the subsequent encoding of selected objects into working memory. In contrast, stored representations of target objects are used during target/nontarget discrimination processes where they are matched with currently selected objects from the search display that are encoded into working memory. In tasks where observers search for dozens of possible target objects (e.g., Wolfe, 2012), this target memory set is too large to be held in working memory and thus has to be represented in a different type of memory store (e.g., activated long-term memory; Cunningham & Wolfe,

2014). This distinction between different types of stores for templates that control attentional selection and representations that are involved in subsequent memory matching processes is less obvious in tasks with only one of two possible target objects, where both types of target representations could in principle be stored in working memory. However, they can still be distinguished with respect to their functions at an early stage of attentional target selection and a later stage of target/nontarget discrimination, respectively.

In summary, the current study shows that the early stages of attentional selectivity are not entirely feature-based, but already involve object-based control mechanisms. Object-based target templates are involved in the guidance of attention towards candidate target objects and in regulating the access of these objects to visual working memory. The effects of these object-based templates emerge later than the effects of feature-based control, but still remarkably early, at around 250 ms after search display onset. However, object-based templates are not employed by default during conjunction search tasks, but only in contexts where feature-based attentional guidance is insufficient to discriminate targets from nontarget objects. The activation of object-based templates may be optional rather than obligatory because the maintenance of such templates in working memory is attentionally demanding. This may also explain why attention can be guided by multiple feature-based templates, but apparently only by a single object-based template at a time. Overall, the current results demonstrate that electrophysiological markers can reveal new insights into the time course of feature-based and object-based attentional control processes in visual search.

Acknowledgement

This work was supported by grant ES/L016400/1 from the Economic and Social Research Council (ESRC), UK. NB is also supported by a New Investigator grant ES/R003459/1 from the ESRC.

References

- Beck, V.M., Hollingworth, A., & Luck, S.J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science, 23*, 887-898.
- Berggren, N., & Eimer, M. (2016a). The control of attentional target selection in a colour/colour conjunction task. *Attention, Perception, & Psychophysics, 78*, 2383-2396.
- Berggren, N., & Eimer, M. (2016b). The guidance of spatial attention during visual search for colour combinations and colour configurations. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1282-1296.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523-547.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: bridging cognition and neurophysiology. *Psychological Review, 112*, 291-328.
- Cunningham, C.A., & Wolfe, J.M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General, 143*, 1585-1599.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.
- Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 578-588.
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology, 99*, 225-234.
- Eimer, M., & Grubert, A. (2014). The gradual emergence of spatially selective target processing in visual search: From feature-specific to object-based attentional control. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1819-1831.
- 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., & Nicholas, S. (2011). What top-down task sets do for us: An ERP study on the benefits of advance preparation in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1758-1766.
- Eimer, M., Kiss, M., Press, C., & Sauter, D. (2009). The roles of feature-specific task set and bottom-up salience in attentional capture: An ERP study. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1316-1328.
- Folk, C.L., & Remington, R.W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 847-858.
- Folk, C.L., Remington, R.W., & Johnston, (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 847-858.
- Grubert, A., & Eimer, M. (2016). All set, indeed! N2pc components reveal simultaneous attentional control settings for multiple target colours. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1215-1230.
- Hickey, C., McDonald, J.J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience*, *18*, 604-613.
- Hopf, J.M., Luck, S.J., Girelli, M., Hagner, T., Mangun, G.R., Scheich, H., & Heinze, H.-J. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex*, *10*, 1233-1241.
- Irons, J.L., Folk, C.L., & Remington, R.W. (2012). All set! Evidence of simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 758-775.
- Jenkins, M., Grubert, A., & Eimer, M. (2017). Target objects defined by a conjunction of colour and shape can be selected independently and in parallel. *Attention, Perception, & Psychophysics*, *79*, 2310-2326.
- Kahneman, D., Treisman, A., & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175-219.
- Jolicoeur, P., Brisson, B., & Robitaille, N. (2008). Dissociation of the N2pc and sustained posterior contralateral negativity in a choice response task. *Brain Research*, *1215*,

160-172.

- Leblanc, É., Prime, D., & Jolicoeur, P. (2008). Tracking the Location of Visuospatial Attention in a Contingent Capture Paradigm. *Journal of Cognitive Neuroscience, 20*, 657-671.
- Lien, M-C., Ruthruff, E., Goodin, Z., & Remington, R.W. (2008). Contingent attentional capture by top-down control settings: Converging evidence from event-related potentials. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 509-530.
- 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.
- Mazza, V., Turatto, M., Umiltà, C., & Eimer, M. (2007). Attentional selection and identification of visual objects are reflected by distinct electrophysiological responses. *Experimental Brain Research, 181*, 531-536.
- McCants, C.W., Berggren, N., & Eimer, M. (in press). The guidance of visual search by shape features and shape configurations. *Journal of Experimental Psychology: Human Perception and Performance*.
- Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology, 35*, 99-115.
- Moore, K.S., & Weissman, D.H. (2010). Involuntary transfer of a top-down attentional set into the focus of attention: Evidence from a contingent attentional capture paradigm. *Attention, Perception, & Psychophysics, 72*, 1495-1509.
- Nako, R., Grubert, A., & Eimer, M. (2016). Category-based guidance of spatial attention during visual search for feature conjunctions. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1571-1586.
- Olivers, C.N.L., Peters, J., Houtkamp, R., & Roelfsema, P.R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences, 15*, 327-334.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology Section A, 40*, 201-237.

- Vogel, E.K., & Machizawa, M.G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*, 748-751.
- Wheeler, M.E., & Treisman, A.M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48-64.
- Wolfe, J.M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*, 202-238.
- Wolfe, J.M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems* (pp. 99-119). New York: Oxford.
- Wolfe, J.M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, *23*, 698-703.
- Woodman, G.F., & Luck, S.J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 121-138.

Figure captions

Figure 1: Example search displays (not to scale) in Experiments 1 and 2. In the double target task of Experiment 1 (upper panel), participants searched for one of two possible targets (in this example, a blue circle or green square). On target-absent trials, displays contained a colour-matching nontarget, a shape-matching nontarget, or an incorrect conjunction object (all appearing in the upper left quadrant in the example shown). In the single target control task, there was only one colour/shape conjunction target, and no incorrect conjunction objects. In Experiment 2 (lower panel), stimuli appeared on the vertical and horizontal midlines, and all had at least one target-matching feature. Targets or incorrect conjunction objects could be in isolation (no-competition trials) or both within the same display (competition trials). Displays could also include two incorrect conjunction objects on target-absent trials.

Figure 2: Reaction times (RTs; line graphs) and error rates (bar charts) for the different task and display conditions in Experiment 1 (left panel; where C-M, S-M, and IC correspond to colour-match, shape-match, and incorrect conjunction conditions respectively) and Experiment 2 (right panel). Error bars represent +/- 1 standard error.

Figure 3: (Left panels) Grand average ERPs obtained in the double target task of Experiment 1 at posterior electrode sites PO7/PO8 contralateral and ipsilateral to a template-matching object. ERPs are shown for the 600 ms interval following search display onset, relative to a 100 ms baseline, separately for targets, incorrect conjunction objects, colour-matching nontargets, and shape-matching nontargets. (Right panel) Difference waveforms computed by subtracting ipsilateral from contralateral ERP, for the four different types of template-matching objects.

Figure 4: Grand average ERPs obtained in the single target control task of Experiment 1 at posterior electrode sites PO7/PO8 contralateral and ipsilateral to targets, colour-matching nontargets, and shape-matching nontargets. ERPs are shown during the 600 ms interval following search display onset, together with the corresponding contralateral-ipsilateral difference waveforms.

Figure 5: (Left panels) Grand average ERPs obtained in Experiment 2 at electrode sites PO7/PO8 contralateral and ipsilateral to targets or to incorrect conjunction objects, shown separately for no-competition (upper row) and competition (lower row) trials. ERPs are shown for the 600 ms interval following search display onset, relative to a 100 ms baseline. (Right panels) Contralateral-ipsilateral difference waveforms for targets and incorrect conjunction objects, shown separately for no-competition and competition trials.

Figure 6: Difference waveforms in Experiment 2 computed by subtracting ipsilateral from contralateral ERPs, separately for targets on trials with fast or slow RTs, and incorrect conjunction objects, for the 600 ms interval following search display onset (collapsed across no-competition and competition trials).

Figure 1

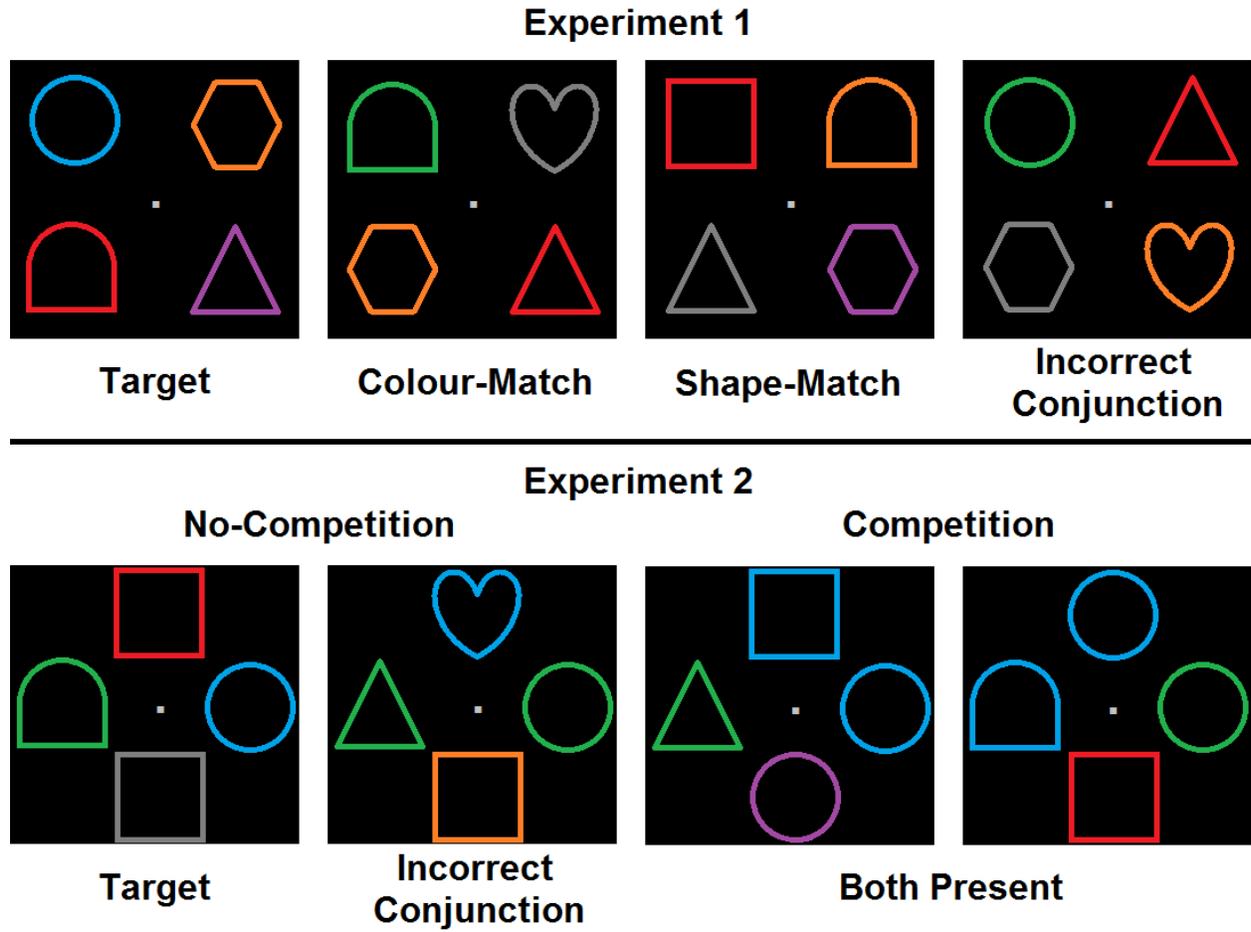


Figure 2

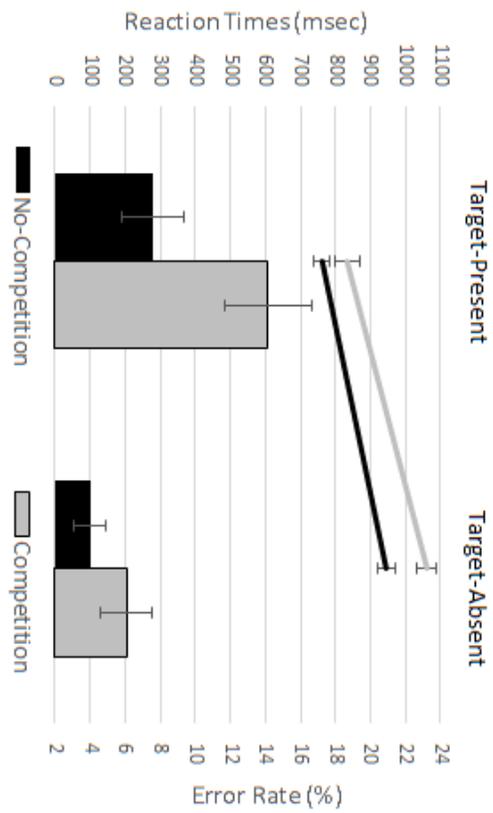
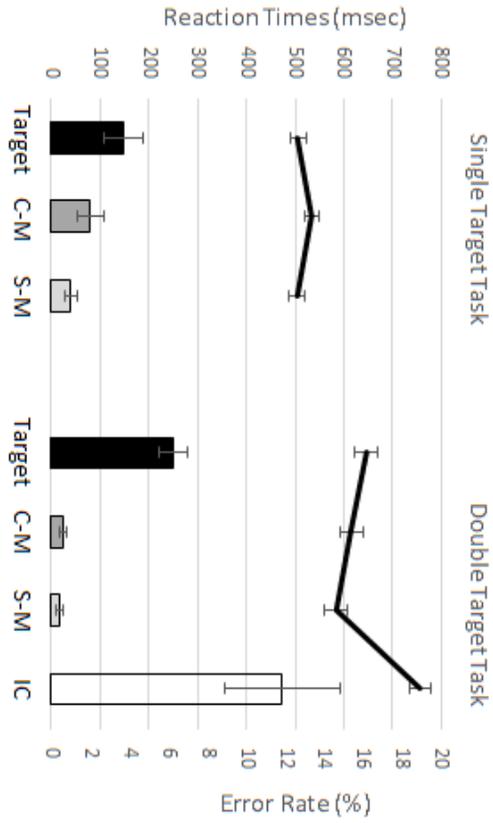


Figure 3

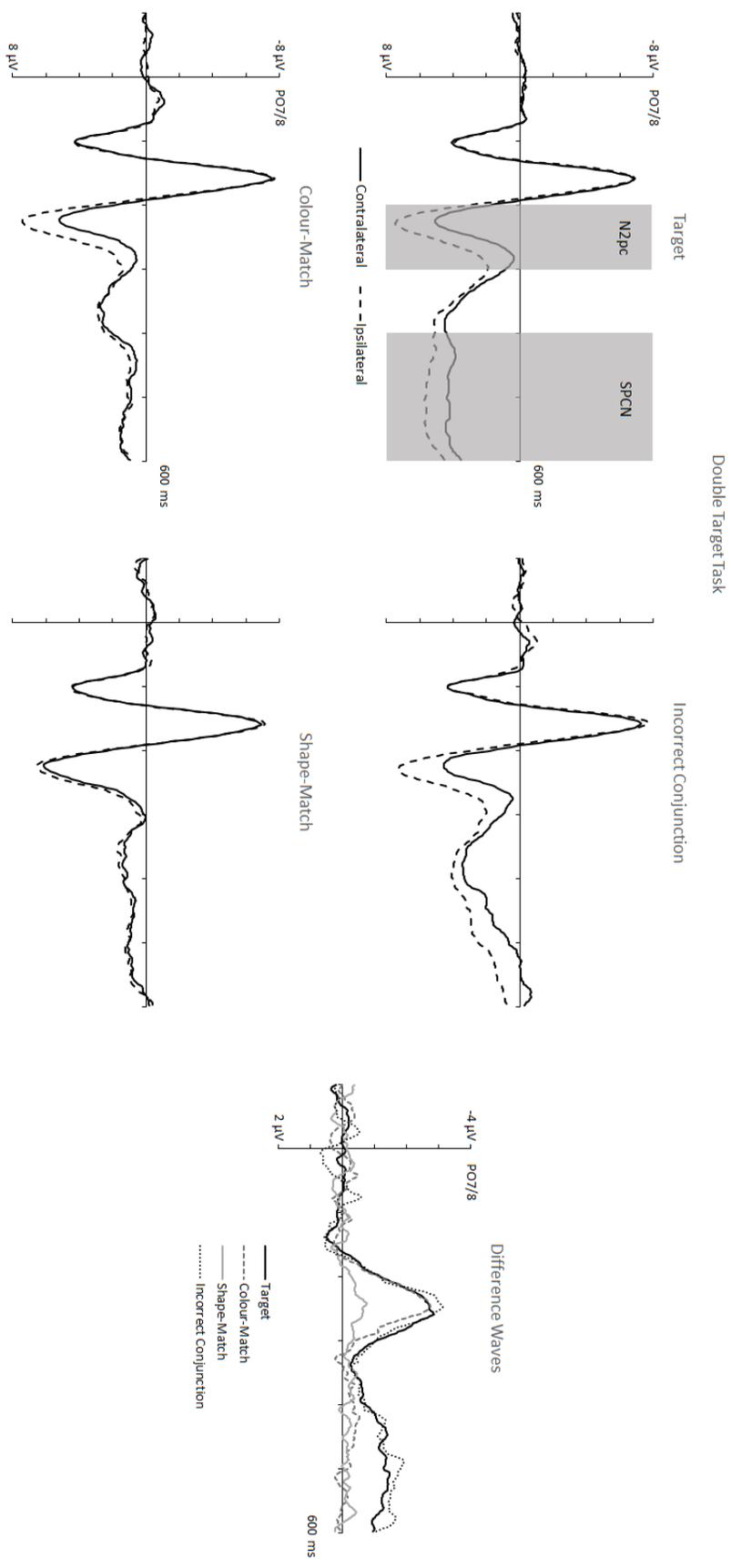


Figure 4

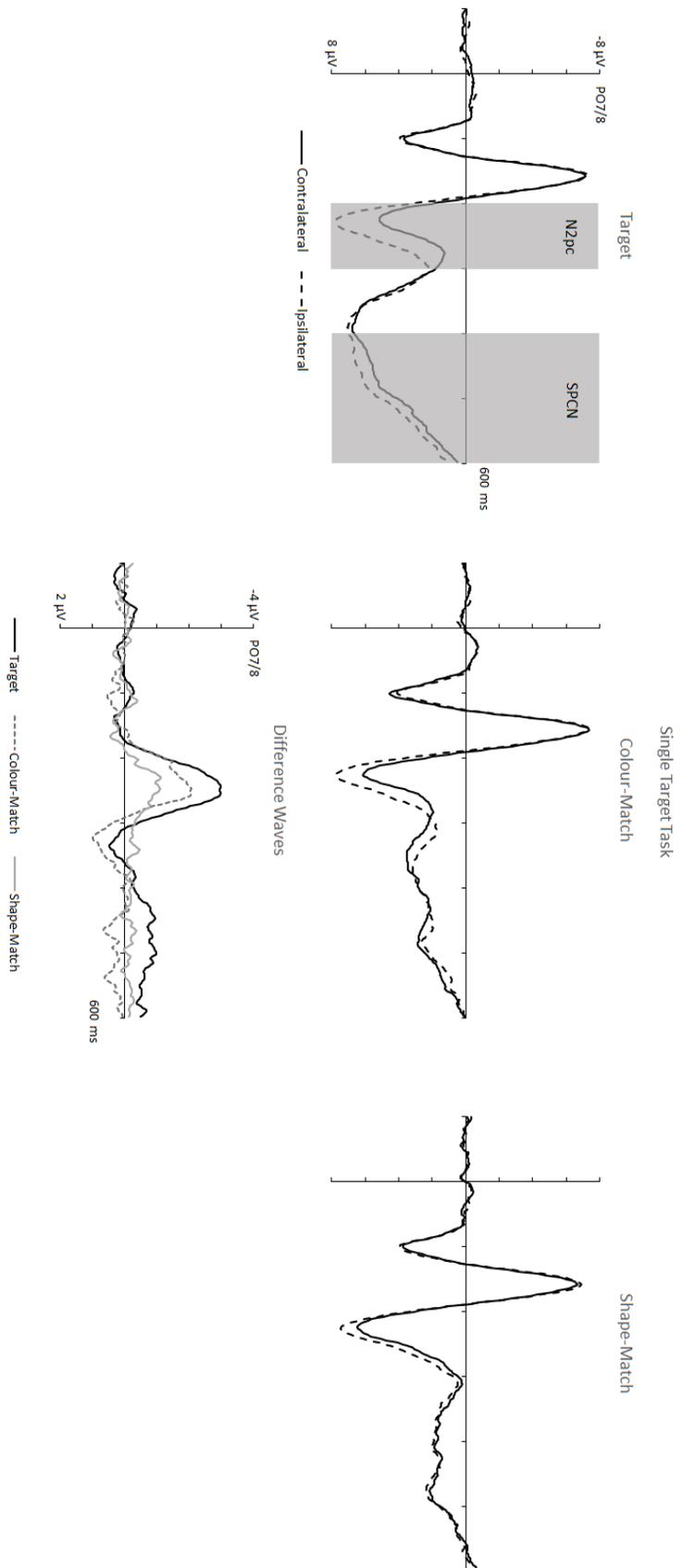


Figure 5

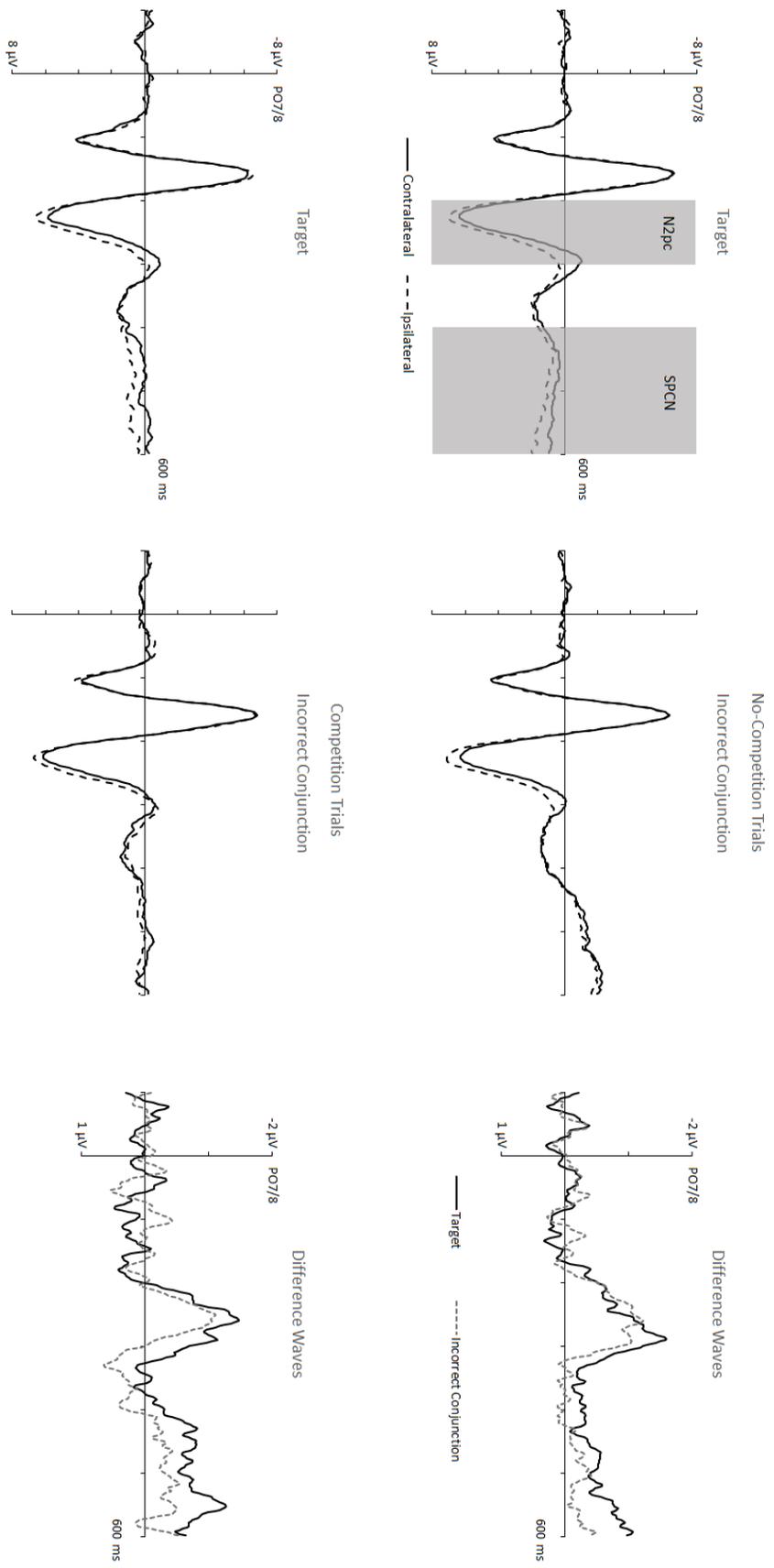


Figure 6

