Research report

The countervailing forces of binding and selection in vision

Adam C. Snyder\textsuperscript{a,b} and John J. Foxe\textsuperscript{a,b,*}

\textsuperscript{a}The Cognitive Neurophysiology Laboratory, Children's Evaluation and Rehabilitation Center (CERC), Departments of Pediatrics and Neuroscience, Albert Einstein College of Medicine, Van Etten Building – Wing 1C, Bronx, NY, USA

\textsuperscript{b}The Cognitive Neurophysiology Laboratory, Program in Cognitive Neuroscience, Departments of Psychology and Biology, City College of the City University of New York, NY, USA

Abstract

Evidence indicates that when one feature of an object is specifically attended, other task-irrelevant features of that object also receive enhanced processing, presumably as a result of automatic binding processes. On the other hand, evidence also shows that attention can be selectively biased in favor of processing one feature at the expense of processing others. Thus, binding invokes combinatorial processing of related features whereas selective attention emphasizes differential processing. We hypothesized that binding and selective attention-based attention depend on a common resource and therefore might show interference effects. The current study tested this by manipulating binding demands while human participants directed their attention to the color or motion of moving random dot stimuli. Response time measures showed that effects of biased attention were reduced when binding demands were increased. This finding supports the thesis that binding relies, at least in part, upon the same mechanisms invoked by preparatory biasing of selective attention.

1. Introduction

Individuals with focal cortical lesions frequently present with highly-specific impairments of processing for one type of stimulus feature while other perceptual processes remain largely unaffected (e.g., Tamler, 1970; Vaina, 1994). These observations highlight a fundamental property of the brain's operation: the various forms of information available in complex environmental signals are processed in parallel channels that rely on distinct neural circuitry (see also Desimone and Ungerleider, 1986; Livingstone and Hubel, 1988; Tanaka, 1996). Information from these separate processing streams is then synthesized to engender a coherent experience characterized by the perception of discrete objects defined by conjunctions of many features. This process of binding together features also has some anatomical specificity, as evidenced by the striking difficulties of some individuals with parietal lesions. These patients have a spared ability to detect the presence of single features, but show impaired ability to identify particular conjunctions of features (Eglin et al., 1989, 1991, 1994; Esterman et al., 2000; Laeng et al., 2002; Pavlovskaya

* Corresponding author. The Cognitive Neurophysiology Laboratory, Children’s Evaluation and Rehabilitation Center (CERC), Department of Pediatrics, Albert Einstein College of Medicine, Van Etten Building – Wing 1C, 1225 Morris Park Avenue, Bronx, NY 10461, USA.
E-mail address: john.foxe@einstein.yu.edu (J.J. Foxe).

0010-9452/$ – see front matter © 2011 Elsevier Srl. All rights reserved.
http://dx.doi.org/10.1016/j.cortex.2011.05.003

http://www.sciencedirect.com
et al., 2002). Often this impaired binding is accompanied by impaired spatial processing and spatial attention, supporting the view that spatial attention is an important component of binding processes (cf. Treisman and Gelade, 1980). Indeed, stressing spatial attention demands increases the occurrence of illusory feature conjunctions in healthy research participants (Treisman and Schmidt, 1982).

In addition to their dependence on spatial attention, binding processes have been known to interact with feature-based attention. It has often been observed that when attention is directed to one feature of an object there is a tendency for task-irrelevant features of that object to receive enhanced processing also (O’Craven et al., 1999; Schoenfeld et al., 2003; Wylie et al., 2004; Molholm et al., 2007; Katzner et al., 2009; Fiebelkorn et al., 2010a, 2010b). The evidence therefore supports a hierarchical model of these processes whereby spatial attention precedes binding, which itself leads to a spread of feature-based attention.

Recently, our group performed the first investigation of the role of alpha-band oscillations in visual feature-based selective attention (Snyder and Foxe, 2010). This study aimed to expand on prior research in selective attention which had shown that alpha-band power increases index the degree of attentional suppression in visuospatial attention tasks (Worden et al., 2000; Sauseng et al., 2005; Yamagishi et al., 2005; Kelly et al., 2006, 2009, 2010; Thut et al., 2006; Rihs et al., 2007; Gomez-Ramirez et al., 2009), intersensory attention tasks (Foxe et al., 1998; Fu et al., 2001; Gomez-Ramirez et al., 2007), sustained visual attention tasks (Dockree et al., 2007; O’Connell et al., 2009) and audiospatial attention tasks (Kerlin et al., 2010; Banerjee et al., in press). We asked if alpha-band mediated suppression would be observed in a feature-based selective attention task. In designing that experiment, we became concerned with the interaction of selective attention and binding processes. We reasoned that if spread of feature-based attention was an essential component of object perception, such a spread of attention might reduce the suppressive mechanisms we sought to measure. Thus, we aimed to reduce the tendency to bind the features in our imperative stimuli by using briefly-presented random dot arrays in which each dot moved at a unique rate; a design with no drawback for that study if the underlying assumption (essential cross-feature spread of attention) was false.

Here, we explicitly test this prior assumption that facilitatory attentional selection is essential for binding. To do this, we specifically designed stimuli intended to place attention and binding demands in direct conflict with each other. Whereas prior investigations of the interaction of spatial attention with binding processes have manipulated attention and observed the effects on binding, we took here the converse approach of manipulating binding demands and observing the effects on selective attention. This approach was in line with the causal role that binding processes seem to have with regards to the spread of feature-based attention within objects.

We used a behavioral paradigm with arrays of dots characterized by one of two types of motion. In one case, all dots moved at the same speed and in the same direction. This type of motion strongly invokes the Gestalt principle of “common fate” (Wertheimer, 1923), leading to a percept of the dots as painted on a moving transparent surface. We termed this condition the STRONG object condition. For the other class of stimuli, all dots moved in the same direction, but each dot had a unique speed. This condition we termed the WEAK object condition, since the percept of a single surface is destroyed when the dots move independently. Our central thesis was that STRONG stimuli would invoke greater binding processes for object perception, and that such binding would counteract suppressive preparatory attention, as evidenced by response time (RT) measures. While increased binding due to Gestalt grouping has not been explicitly demonstrated, we deemed this was a reasonable assumption due to the clearer object-nature of the STRONG stimulus.

This paradigm was a variant of the classic cued-attention task. In a cued-attention task, a symbolic cue (“stimulus one”: S1) at the beginning of each trial provides task-relevant information about the nature of an upcoming imperative stimulus (S2). While the cue information is often helpful, or valid, on occasion the cue is misleading, or invalid. Non-informative neutral cues are included as a basis for comparison. For this experiment, cues directed attention to either the color or motion of the S2, and participants had to make a discrimination regarding one of those features. A typical finding is that valid cues lead to faster RTs compared to neutral cues whereas invalid cues lead to slower RTs compared to neutral cues (Posner, 1980). This pattern of results is taken to indicate that participants are able to utilize the cue information to prepare attention to facilitate the processing of the information that the cue had indicated while suppressing the processing of competing information.

In this case, the status of the imperative stimulus as an object was irrelevant to the task. Thus, if attentional selection was non-essential for binding in the face of a STRONG stimulus, then the optimal strategy would be to maintain suppression of irrelevant features. However, if attentional selection was essential for binding, then the presence of a STRONG stimulus would reduce suppressive attention effects compared to those with respect to a WEAK stimulus. We used RT as a dependent measure to test these predictions.

Specifically regarding the hypothesis that attention was essential for binding, we predicted the following:

1) For WEAK stimuli, valid cues would result in speeded RTs compared to neutral cues (RT benefit) and invalid cues would result in delayed responses (RT cost).
2) Compared to WEAK stimuli, RT benefits for STRONG stimuli preceded by valid cues would be reduced. This is because the anticipatory attentional suppression of the irrelevant feature would have to be counteracted to support binding for object perception.
3) Compared to WEAK stimuli, RT costs for STRONG stimuli preceded by invalid cues would be unchanged. This is because the anticipatory state would be equally biased against processing the invalidly-cued target feature regardless of the stimulus’ status as an object (i.e., STRONG or WEAK).
2. Method

2.1. Participants

Sixteen adults (8 males, 3 left-handed) aged 20–47 years (mean ± SD: 27.58 ± 7.27) participated in the experiment. Participants were sourced from the undergraduate and graduate student populations of The City College of New York, and from the local community. None of the participants had any history of brain injury or disease, per self-report. Participants had normal or corrected-to-normal visual acuity per self-report. Fifteen participants had normal color vision. One participant was unable to discriminate red and green hues, but was able to perform the task normally using blue/yellow discrimination. All participants provided informed consent prior to the experiment. All materials and procedures were approved by the institutional review board of The City College of New York in accordance with the United States Public Health Service Act (US 45 CFR 46).

2.2. Experimental design

Each participant completed two experiments involving judgments of random dot kinematograms. In one experiment, each dot in the stimulus moved at a unique speed (WEAK condition). In the other experiment, all dots in the stimulus moved at the same speed (STRONG condition). Experiments were administered in a counter-balanced order across participants. Participants completed the experiments in two separate sessions divided by an interval of at least two weeks. The purpose of this separation was to reduce task-set interference effects. Due to attrition, four participants completed only the WEAK condition and three participants completed only the STRONG condition. The remaining nine participants completed both conditions.

2.3. Task

Procedures were adapted from a prior feature-based attention experiment (Snyder and Foxe, 2010). We used a cued-attention paradigm in which visual word forms (‘COLOR’, ‘HUE’, ‘MOTION’, or ‘DIRECTION’) indicated to the participant the relevant feature to attend. A non-informative neutral cue (‘NOTHING’) was used for comparison. The words ‘color’ and ‘hue’ both directed attention to the color of the stimulus. Likewise, ‘motion’ and ‘direction’ both directed attention to the motion of the stimulus. The use of multiple cue words for each feature was to reduce the automatization of the task through implicit learning, with an aim to maintain the engagement of endogenous orienting mechanisms throughout the session. Participants were instructed to use the cue information to guide them in detecting a target in a subsequent imperative stimulus (S2), and upon detection of a target to respond with a button press as quickly as possible without sacrificing accuracy. Targets were present on 50% of trials, and were characterized by a particular property in either feature dimension (Fig. 1b). Twenty percent of targets were preceded by a neutral cue. Of informatively-cued targets, 80% were preceded by a valid cue (the cued feature was the feature that defined the target), and the remaining 20% were preceded by an invalid cue (the target was defined by the uncued feature). Participants were instructed to respond to all targets, including those that were invalidly cued.

2.4. Stimuli

The experiment was administered in a light- and sound-attenuated chamber using Presentation software version 14.4 (Neurobehavioral Systems). All stimuli were presented on a standard size cathode ray tube (CRT) monitor with a 60 Hz refresh rate. Trials began with a warning cue consisting of a white fixation dot on a black background for 1 sec, followed by a cue word in white block capitals for 1 sec. After an interval of 1.7–2.3 sec (random, and evenly distributed) during which only a black screen was displayed, the S2 was presented for .2 sec. Each S2 was followed by a 1 sec response interval. Each subsequent trial began immediately following the response interval (Fig. 1a).

The S2 consisted of an array of one thousand dots, each subtending .05° of visual angle, constrained to a square aperture subtending 5° of visual angle. In the WEAK condition, each dot moved at a unique speed between 14° and 28° of visual angle per second (evenly distributed). In the STRONG condition, all dots moved at a speed of 21° of visual angle per second. Dots ‘wrapped around’ the edges of the square aperture, so that the total amount of illumination was held constant.

Dots were colored with hues from an isoluminant plane of DLK color-space (Derrington et al., 1984). This color-space uses the response properties of neurons in macaque lateral geniculate nucleus to create a subjective luminance axis, planes orthogonal to which are approximately isoluminant. The use of this color-space enables the continuous variation of hue needed to derive hue discrimination thresholds while controlling for subjective luminance.

2.5. Targets

For both WEAK and STRONG conditions, standard trials were characterized by all dots moving on a common linear trajectory and having the same color. For color targets in both the WEAK and STRONG conditions, 20% of dots had a different color than the majority while the dots continued to move as on standard trials for the respective condition (i.e., WEAK or STRONG). For motion targets in the WEAK condition, 20% of the dots moved on a different linear trajectory than the majority. For motion targets in the STRONG condition, all dots moved on a common curved trajectory. Schematized examples of targets are illustrated in Fig. 1b. The degree of difference for each of the relevant targets was titrated on a per-subject basis to 80% detection rate prior to beginning each experiment using an up–down transformed response (UDTR) modified staircase procedure (Wetherill and Levitt, 1965). No particular value of any feature indicated a target: subjects had to detect a particular feature variation in the stimulus. This strategy was used to reduce competition within a feature processing area (if subjects were attending to red and suppressing green, for example). The goal, rather, was to have subjects attend to color and suppress motion, or vice versa as the cue indicated.
2.6. Analysis

RT data were analyzed using SPSS Statistics software version 17.0. We used a $3 \times 2 \times 2$ univariate analysis of variance (ANOVA) with random factors of ‘validity’ (three levels: invalid, neutral and valid), ‘feature’ (two levels: color and motion) and ‘object’ (two levels: STRONG and WEAK). The factor of ‘subject’ was included as a covariate in the model to account for overall differences in RT between subjects, which were not an effect of interest. Rather, we were interested in relative effects within subjects due to the parameters of the cue and S2. Only participants that completed both experiments were included in the ANOVA ($N = 9$). All correct responses (‘hits’) were included in the analysis for a total of 7395 observations (4308 hits in the WEAK condition and 3087 hits in the STRONG condition) from a set of 18,122 total trials (10,613 in the WEAK condition and 7509 in the STRONG condition). An alpha criterion of $p \leq .05$ was used to define significance.

3. Results

Group-level RT data are illustrated in Fig. 2. A main effect was found only for the factor of ‘subject’ [$F_{(1,7382)} = 9.425, p < .01$], which was a nuisance variable in this analysis. There was a trend toward significance on the factor of ‘validity’ [$F_{(1,7382)} = 5.072, p = .089$]. Overall, RTs following valid cues (629.2 ± 3.1 msec, mean ± SEM) were faster than those following neutral cues (653.9 ± 5.5 msec), whereas RTs following invalid cues (694.4 ± 6.2 msec) were slower than those following neutral cues. RTs did not differ between STRONG and WEAK stimuli [674.1 ± 4.5 msec and 644.2 ± 3.8 msec, respectively; $F_{(1,7382)} = 1.339, p = .411$]. RTs did not differ between color and motion targets [657.8 ± 4.2 msec and 660 ± 4.2 msec, respectively; $F_{(1,7382)} = .010, p = .932$], consistent with our aim to match task difficulty across the two features.
The ‘validity’ × ‘object’ interaction was significant \(F_{(2,7382)} = 18.945, p = .050\). RTs were faster following valid cues than neutral cues, but only when S2s were WEAK (605.2 ± 4.0 msec and 641.6 ± 7.2 msec for valid and neutral cues, respectively). When S2s were STRONG, RTs following valid cues (653.2 ± 4.7 msec) did not differ from those following neutral cues (666.1 ± 8.4 msec). RTs to validly-cued STRONG S2s were significantly slower than those to validly-cued WEAK S2s. RTs following invalid cues were slower than those following neutral cues for both STRONG (invalid: 703.1 ± 9.4 msec) and WEAK (invalid: 685.7 ± 8.0 msec) S2s.

The ‘validity’ × ‘feature’ interaction was significant \(F_{(2,7382)} = 31.893, p = .031\). RTs to validly-cued color targets (620.1 ± 4.4 msec) were faster than those to validly-cued motion targets (638.3 ± 4.2 msec).

The ‘feature’ × ‘object’ interaction was significant \(F_{(1,7382)} = 59.827, p < .01\). RTs to WEAK motion targets (633.6 ± 5.2 msec) were faster than those to WEAK color targets (654.8 ± 5.6 msec), whereas RTs to STRONG motion targets (687.4 ± 6.5 msec) were slower than those to STRONG color targets (660.9 ± 6.2 msec).

The three-way interaction (‘validity’ × ‘feature’ × ‘object’) was not significant \(F_{(2,7382)} = .184, p = .832\).

4. Discussion

We found that attention could be successfully biased between the visual features of color and motion as indexed by RT effects. That is, the classic cost–benefit effects of invalidly versus validly cuing an upcoming to-be-attended feature dimension were observed here, with participants showing slowed reaction times to invalidly-cued targets and speeded reaction times when the cue accurately predicted the feature dimension of the impending judgment. However, this ability to selectively attend was modulated by the motion type of the S2, which was specifically manipulated to differentially engage binding processes. The key finding here is that the deployment of anticipatory feature-based attention was considerably more effective when the invocation of automatic binding processes was weak.

To recall, we made three specific predictions, each of which was confirmed. Firstly, we predicted that RTs for validly-cued WEAK stimuli would be faster than RTs for neutrally-cued WEAK stimuli (i.e., a RT benefit for valid cues), which was indeed the case. Secondly, we predicted that STRONG stimuli would show a smaller RT benefit. In fact, STRONG stimuli did not show a benefit for valid cues at all. Finally, we predicted that RT costs for invalid cues would not differ between WEAK and STRONG conditions, because attention was equivalently biased against the target feature in each case. This prediction was also confirmed.

4.1. Physiological implications

The interference effects illustrated by this study suggest that feature-based selective attention and binding for object perception are processes that share at least one common anatomical substrate or physiological mechanism. One candidate mechanism that could be common to both of these processes is alpha-band oscillations. Changes in alpha-band power have been observed for manipulations of both feature-based attention and object perception.

We have previously reported a role for alpha-band oscillations in feature-based selective attention (Snyder and Foxe, 2010). Specifically, alpha-band power increased during preparatory intervals in brain regions linked to the processing of the potentially distracting feature, a pattern of results consistent with the suppressive role of alpha-band power increases reported in spatial selective attention studies (e.g., Worden et al., 2000; Kelly et al., 2006; Thut et al., 2006).

In an MEG study of object-processing, Vanni et al. (1997) reported greater decreases in alpha-band power following the presentation of intact line drawings compared to their scrambled and incoherent counterparts. The researchers linked this alpha-band effect to object-perceptual processes (as opposed to low-level stimulus effects) by asking subjects to report after each trial whether they perceived a coherent object or not. Short (30–106 msec) durations were used such that subjects typically only correctly reported perceiving coherent objects 50% of the time. Alpha-band responses to misidentified objects were intermediate to those of correctly identified objects and correctly identified scrambled images, indicating that alpha-band decreases relate to the perception of the stimulus as an object, rather than its stimulus attributes.

Alpha-band oscillations have therefore been implicated in both feature-based selective attention and coherent object perception, making them a promising candidate for further study of the physiological substrates underlying the interplay between attention and object-perception processes. It is noteworthy that alpha-band activity emanates predominantly from the right parietal lobe; precisely the regions implicated in both attentional deployment (Heilman and Van Den Abell, 1980; Vallar and Perani, 1986; Corbetta et al., 2000; Foxe et al., 2003) and in the attentional coordination of binding (Eglin et al., 1989, 1991, 1994; Esterman et al., 2000; Laeng et al., 2002; Pavlovskaya et al., 2002). Other potential avenues of investigation include other oscillatory frequency bands. In particular, object-perceptual processes have been linked to the beta (sehatpour et al., 2008) and gamma (eckhorn et al., 1988; gray et al., 1989; tallon-baudry et al., 1996) bands.

4.2. Alternative interpretations

Because we administered STRONG and WEAK stimuli in separate sessions, it is possible that participants’ knowledge of the nature of the stimuli influenced their preparatory approach. If participants did not bias attention to the same extent in anticipation of a STRONG stimulus as they did in anticipating a WEAK one, this could result in decreased RT benefits for valid cues. However, this would also be expected to result in decreased RT costs for invalid cues, which was clearly not what was found. This suggests that the preparatory approach was the same for both classes of stimuli, and that decreased biasing due to binding demands manifests during stimulus-processing stages.

From a theoretical standpoint, it is unlikely that the quality of processing of the relevant feature is reduced due to...
increased binding demands. Indeed, this would be contrary to the literature regarding the spread of attention within an object, which has hitherto reported only enhancement effects (e.g., O’Craven et al., 1999; Schoenfeld et al., 2003; Wylie et al., 2004; Molholm et al., 2007; Katzner et al., 2009; Fiebelkorn et al., 2010a, 2010b). Furthermore, if the quality of processing of the relevant feature was reduced, RTs following neutral cues would be predicted to be slower for STRONG stimuli compared to WEAK stimuli, and this was not observed here. Thus we conclude that the quality of processing of the irrelevant stimulus feature was likely enhanced during STRONG object perception, resulting in a decreased biasing between the two competing features. The key implication of this is that a certain “quality threshold” of processing must be met for features to be bound together, and that this requirement dominates prior attentional deployments during object perception.

4.3. Methodological considerations

A critical feature of the current study is that participants’ performance levels were held below ceiling, indicating that processing capacity was exhausted. Tasks performed at ceiling levels may leave processing resources to spare such that the functions under investigation do not have to compete, obviating the need for attentional biasing. That is, it may be the case that enhanced processing of irrelevant features in a selected object is a convenient thing to do with excess processing resources, but is not strictly a fundamental part of the object-perception process.

For example, O’Craven et al. (1999), in a neuroimaging study, showed participants overlaid images of faces and houses, one of which moved on each trial. In the critical condition, the participants attended to the direction of motion of the moving image, but the content of that image was irrelevant to the task. The key result was that metabolic signatures increased in face-responsive brain areas when the face was moving and increased in house-responsive areas when the house was moving, notwithstanding the task-irrelevance of these object categories. This result was taken to indicate the spread of attention to the task-irrelevant category of the attended image. However, in this case the task was likely quite easy (behavioral results were not reported, but participants merely had to discriminate between ‘up’, ‘down’, ‘left’, and ‘right’), and there would be no disadvantage conferred by allocating excess attentional resources to processing the object category. In fact, it may be the case that parsing the object category assisted the participants in segregating the moving image from the spatially colocalized static image. Thus, one cannot conclude from these results that the spread of attention to irrelevant features is a fundamental process rather than a merely convenient one.

To test the robustness of the cross-feature attentional-spreading phenomenon, it is necessary to create a context for which it would be advantageous for attention not to spread between the features of an object (i.e., the two features are competitors for the same processing resources) as in the current study. In such a context, evidence of attentional spread (such as decreased biasing of attention between the shared features of an object) would indicate the automaticity of the process. To our knowledge, the current study is the first to demonstrate that cross-feature spread of attention occurs even when explicitly detrimental to task performance.

4.4. Other findings

In addition to the predicted interaction between cue validity and object strength, we found two-way interactions between which feature was cued and the effects of cue validity and object strength. These interactions could be related to the temporally integrative nature of motion judgments. For example, we found that RTs to validly-cued motion targets were delayed 18 msec compared to validly-cued color targets. This delay might reflect the fact that motion direction required at least two video frames to determine, whereas color information was available in a single frame. Likewise, STRONG motion targets, which had curved motion, required at least three frames to resolve (i.e., two video frames to establish direction, and then an additional frame for the change in direction). It would not be surprising if the brain required more than this minimum number of frames to assess the motion curvature in this demanding task. This may account for why this condition had the greatest delay in RT compared to other target types.

4.5. Conclusion

In conclusion, we found that RT benefits for valid feature-based attentional cues that were observed when imperative stimuli only weakly engaged binding processes were reduced when imperative stimuli strongly engaged binding. This result suggests that feature-based selective attention processes and binding processes for coherent object perception rely at least partially on a common physiological substrate. One potential candidate substrate is alpha-band oscillations, which have been separately implicated in both processes.

Acknowledgments

This work was supported by a grant from the U.S. National Science Foundation to Professor Foxe (BCS0642584). Mr. Snyder is supported by a Ruth L. Kirschstein National Research Service Award (NRSA) pre-doctoral fellowship from the National Institute of Mental Health (NIMH – F31 – MH087077).

References

Banerjee S, Snyder AC, Molholm S, and Foxe JJ. Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual target locations: Supramodal or sensory-specific control mechanisms? *Journal of Neuroscience*, in press.


Dockree PM, Kelly SP, Fuxe JI, Reilly RB, and Robertson IH. Optimal sustained attention is linked to the spectral content of background EEG activity: Greater ongoing tonic alpha (approximately 10 Hz) power supports successful phasic goal activation. European Journal of Neuroscience, 25(5): 900–907, 2007.


