VISUAL COGNITION, 2010, 18 (10), 1486–1541

Psychology Press
Taylor & Francis Group

Object Perception, Attention, and Memory (OPAM) 2010 Conference Report
18th Annual Meeting, St. Louis, Missouri, MO, USA

Organized by Michi Matsukura, Melanie Palomares,
Brian Levinthal, and Timothy Vickery

---

http://www.psypress.com/viscog DOI: 10.1080/13506285.2010.521658

Testing capacity limitations of surface completion using the simultaneous-sequential method

Mouna Attarha and Cathleen M. Moore
University of Iowa, Iowa City, IA, USA

Natural scenes are comprised of objects and surfaces in three-dimensions, and it is unusual to see a single surface in isolation. Instead, we tend to see multiple surfaces at different depths, with closer surfaces partially blocking regions of farther surfaces. This produces absences of local stimulus correlates in the retinal image. Nonetheless, the visual system represents complete surfaces. We use the term *surface completion* to refer to the set of processes that give rise to the perception of complete rather than fragmented surfaces.

Despite the apparently fundamental nature of surface representation in vision (Gibson, 1979; Marr, 1982), surface completion in many classic theories of visual perception is either not considered, or tends to be associated with relatively late processing in association with object recognition (e.g., Biederman, 1987; Treisman & Gelade, 1980; Wolfe, 1994). According to these theories, image features, such as colour, orientation, and luminance, are filtered more-or-less passively from the retinal image, and are hierarchically combined to allow for higher level cognitive functions, such as object recognition. In contrast to feature-focused theories, Nakayama and colleagues (e.g., Nakayama, He, & Shimojo, 1995) have argued that the retinal image is first organized into apparent surfaces (based on image features), and that prior to this organization, there is no conscious access to visual information. According to this view, the filter characteristics of neurons in anatomically early visual cortex serve, at least initially, to establish surface representations. Conscious access to feature information, such as in visual search and texture segregation, involves only indirect access to the

---

Please address all correspondence to Mouna Attarha, 103 Spence Laboratories of Psychology, Department of Psychology, University of Iowa, Iowa City, IA 52242, USA. E-mail: mouna-attarha@uiowa.edu

output of those units, mediated through surface representations. In support of this view, Nakayama and colleagues conducted experiments in which they manipulated surface organization via binocular disparity while introducing virtually no change to the image features in the display. They found effects of the surface organization on perception, visual search, and the cued allocation of attention, despite there being no changes in the feature information.

An implication of the surfaces-first view of visual processing is that the establishment of surface organization should depend only on unlimited-capacity processes. That is, the completion of surfaces within a scene should not depend on whether there are other surfaces in the scene. We applied the simultaneous-sequential method (e.g., Eriksen & Spencer, 1969; Scharff, Palmer, & Moore, in press) to ask whether surface completion involves limited- or unlimited-capacity processes.

We examined capacity limitations of modal completion (i.e., “illusory surfaces”) and amodal completion (i.e., completion of surfaces behind occluding surfaces), as illustrated in Figure 1A. Observers were shown four stimuli, one of which was a target. The target was a horizontal or a vertical surface, which was created by rotating the inducing “pacmen” (Ringach & Shapley, 1996). The other three (distractors) were square surfaces. The task was to find and report which one of the two targets was in the display. The stimuli were presented either all at once (simultaneous condition) or in two sets of two sequentially (sequential condition). The target and distractors were available for exactly the same amount of time across these conditions (see Figure 1B). If surface completion engages only unlimited-capacity processes, then accuracy should be as high in the simultaneous condition as in the sequential condition, because the completion of the different surfaces should be unaffected by having to complete other surfaces. Alternatively, if surface completion engages limited-capacity processes, then there should be an advantage for the sequential condition over the simultaneous condition, because there are fewer surfaces to interfere with each other at any one time in the Sequential condition. Following Scharff et al. (in press), we also included a repeated condition in which all four stimuli were presented simultaneously twice, thereby providing two opportunities to find the target. This condition provides a baseline for testing among specific limited-capacity models, and allowed us to confirm that there was room for improvement in the event that a null result was found between simultaneous and sequential conditions.

METHOD

Five observers were tested in two experiments for eight sessions each (two practice and six “real”). Stimuli and trial-event details are illustrated

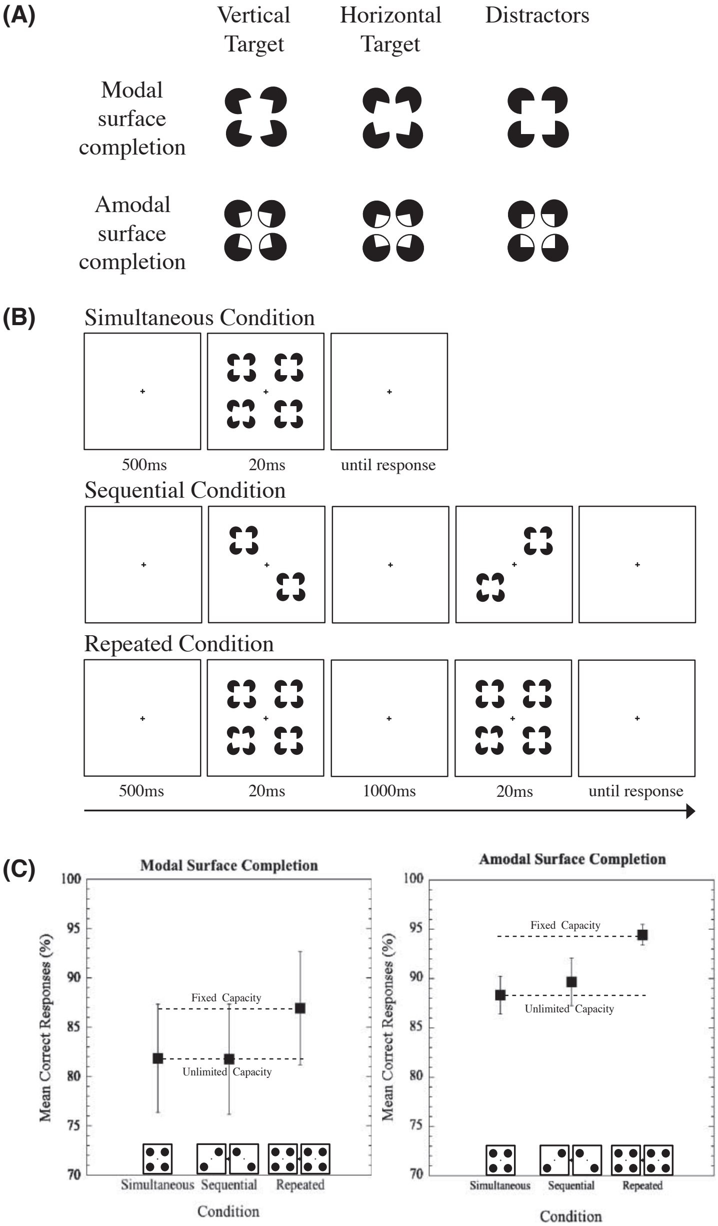


Figure 1. (A) Modal and amodal targets and distractors. (B) Trial events for simultaneous, sequential, and repeated conditions. (C) Mean correct responses (%) as a function of condition collapsed across observers for modal and amodal surface completion.

in Figure 1A–B. Display condition was randomly mixed within blocks of trials. Target type (horizontal, vertical) and target location were random.

RESULTS

The left panel of Figure 1C shows the results from Experiment 1 (modal completion). The data are consistent with unlimited-capacity processing. Accuracy was not significantly different between the simultaneous (82%) and sequential (81%) conditions, $t(4) = -0.74$, *ns*, and accuracy in the repeated condition (84%) was greater than the average of the other two, $t(4) = 5.26$, $p < .01$. All five observers showed this pattern.

The right panel of Figure 1C shows the results from Experiment 2 (amodal completion). The data from this experiment are also consistent with unlimited-capacity processing. Accuracy was again not significantly different between the simultaneous (88%) and sequential (90%) conditions, $t(4) = 1.66$, *ns*, and accuracy in the repeated condition (94%) was greater than the average of the other two, $t(4) = 4.48$, $p < .01$. Three out of the five observers showed this pattern.

IMPLICATIONS

The hypothesis that surface representation is a fundamental early visual processing stage implies that it unfolds in the same way, regardless of how many surfaces are being completed (i.e., it engages only unlimited-capacity processes). This conclusion would be strengthened if we had tested a greater number of surfaces. This presents a challenge, however, because adding stimuli to increase the number of surfaces in these displays introduces the potential of interstimulus interference unrelated to surface completion itself (e.g., crowding). At present, by applying the simultaneous-sequential paradigm to modal and amodal surface completion, we found evidence consistent with the surfaces-first view for up to at least four surfaces.

REFERENCES

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology*, *79*, 1–16.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W. H. Freeman.

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *Vision: In invitation to cognitive science* (pp. 1–70). Cambridge, MA: MIT Press.

Ringach, D., & Shapley, R. (1996). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision Research*, *36*, 3037–3050.
Scharff, A., Palmer, J. P., & Moore, C. M. (in press). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance*.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*, 202–238.

Whatever you do, don't look at the . . .: Exploring the parameters of an exclusionary attentional template

Valerie M. Beck, Steven J. Luck, and Andrew Hollingworth
University of Iowa, Iowa City, IA, USA

Previous research has demonstrated that observers can use feature values stored in working memory to form an "attentional template" and guide search towards items that match those feature values. We have previously demonstrated that observers are quite efficient at restricting search on the basis of one feature value (Beck & Luck, 2009) and are even able, though less efficient, to restrict search on the basis of two feature values (Beck & Luck, 2010). In the current task, we tested whether observers would be able to form an exclusionary attentional template to bias search away from, and thus avoid searching, items that matched a particular, irrelevant colour.

The current task was to locate a target item (circle with a gap in the top or bottom) amongst distractor items (circles with a gap in the left or right) while gaze was tracked by an SR Research EyeLink 1000 system. Observers responded manually to indicate the location of the gap in the target. Each array contained 16 items (one target, 15 distractors), evenly divided into either four colours (four items of each colour) or eight colours (two items of each colour). Each search array was preceded by a cue (coloured square) that appeared briefly (see Figure 1a). In one condition, the cue indicated the colour of the target item; in another condition, the cue indicated a colour that the target item would *not* be; and in a third condition, the cue was an eight-colour checkerboard indicating that the target item could be any colour (neutral cue). The three conditions were blocked and the order was randomized across subjects.

Please address all correspondence to Valerie M. Beck, University of Iowa, Psychology Department, E11 Seashore Hall, Iowa City, IA 52242, USA. E-mail: valerie-beck@uiowa.edu

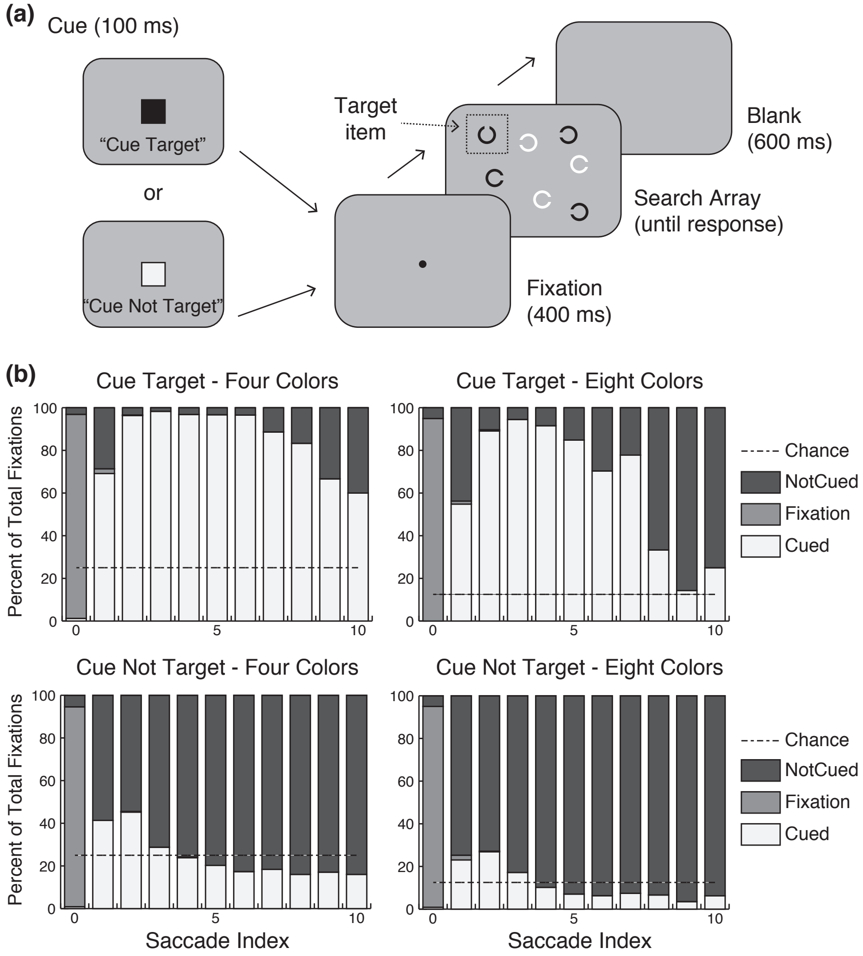


Figure 1. (a) Sample trial: A coloured square presented for 100 ms at the beginning of a trial cued a colour to search, a colour to avoid, or was neutral (not pictured). After a 400 ms fixation screen, the search array (16 items: 4 each of 4 colours or 2 each of 8 colours) appeared and remained on until the observer made a response. Observers were asked to find an item with a top or bottom gap and report the location of the gap. Last, a blank screen was presented until the next trial began (about 600 ms). Conditions ("cue target", "cue not target", and "neutral") were blocked and presented in a random order across subjects. Please note, sample trial is for illustrative purposes only and is not drawn to scale. (b) When the cue indicated the target item colour ("cue target"), observers made saccades to items that matched the cue colour more so than predicted by chance throughout the trial, suggesting they restricted search to relevant items. However, when the cue indicated which colour to avoid ("cue not target"), observers still made saccades to items that matched the cue colour more so than predicted by chance for the first 2–3 saccades, suggesting they were unable to suppress searching items that matched the cue colour until later in the trial.

In accordance with previous studies, observers had no difficulty restricting their search to cued-colour items when the cue indicated the colour of the target item; observers were always more likely to visit an item of the cued-colour (85–90%) than any other colour (1–2% for each colour) throughout the course of a trial (see Figure 1b). On the other hand, observers were not able to avoid cued-colour items in the exclusionary condition until the latter part of a trial. That is, observers often directed gaze to items of the cued-colour (36%/23%) rather than any of the uncued-colours (21%/11% for each colour) during the first few saccades, even though the cued-colour items should be avoided as they were never the target (see Figure 1b). However, once they had some time to exert control (after the first couple eye movements), observers visited the cued-colour items less frequently (14%/6%) and visited the uncued-colour items more frequently (28%/13% for each colour). In the neutral cue condition, observers visited items of each colour roughly equally (25% for each of four colours, and 12.5% for each of eight colours) throughout the course of a trial. These results suggest that it is not possible to effectively use a single colour to form an exclusionary attentional template.

An attentional template helps both to select relevant information and to filter out irrelevant information, but it is not known yet precisely how this is accomplished. In theory, it seems possible to select locations based on a feature value, but then suppress allocation of attention to those locations to filter out irrelevant items. However, the results presented here do not support this possibility, since observers were initially drawn to fixate irrelevant items that matched the cue colour in the exclusionary condition. The pattern of results is consistent, though, with Treisman's Feature Integration Theory of Attention, which suggests that a location has to be attended in order to know what features are present (Treisman & Gelade, 1980). On the other hand, Wolfe's Guided Search Model suggests that priority for searching particular locations can be adjusted on the basis of whether or not those locations contain relevant features (Wolfe, Cave, & Franzel, 1989). If this were the case, we would have expected to see very few fixations to the cued-colour items in the exclusionary condition, but perhaps an increased latency of the first saccade to efficiently suppress those locations. Importantly, though, neither model takes into account the influence of working memory representations on the allocation of attention.

The Biased Competition Theory (Desimone & Duncan, 1995) suggests that working memory representations can bias search towards items that contain shared features. In the current task, observers are cued on a trial-by-trial basis and arguably store a representation of the cue in working memory, which may bias their search towards items matching the cue colour for the first couple saccades. It is possible that, if the cue colour was blocked and observers did not have to rely on a working memory representation, we would see fewer fixations to cued-colour items in the exclusionary condition. Results from the current study suggest that, initially, working memory representations guide attention in terms of enhancing perceptual processing at locations containing the feature value currently active in working memory; these representations cannot be used efficiently to suppress perceptual processing and exclude items from search. This indicates a key limitation on the ability to use working memory representations to guide attention.

REFERENCES

Beck, V. M., & Luc, S. J. (2009, May). *Temporal dynamics of the attentional template during visual search*. Poster presented at the ninth annual meeting of the Vision Sciences Society, Naples, FL.

Beck, V. M., & Luck, S. J. (2010, May). *Guidance of attention during visual search: Can multiple attentional templates operate concurrently?* Poster presented at the 10th annual meeting of the Vision Sciences Society, Naples, FL.

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

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

The attentional blink is attenuated for objects of expertise

Kara J. Blacker and Kim M. Curby
Temple University, Philadelphia, PA, USA

Human information processing capacity is limited at many stages of processing. However, evidence suggests that extensive practice allows individuals to circumvent some capacity limitations. For example, visual short-term memory capacity is greater for objects of expertise than nonexpertise (Curby, Glazek, & Gauthier, 2009). If visual expertise attenuates processing limitations involved in perception, experts might show a reduced attentional blink (AB) for objects from their domain of expertise. The AB refers to an impaired ability to process the second of two targets embedded within a rapid, sequentially presented stream of items (Broadbent & Broadbent, 1987; Raymond, Shapiro, & Arnell, 1992). One

---

Please address all correspondence to Kara Blacker, Temple University, 1701 N. 13th St., Philadelphia, PA 19122, USA. E-mail: kara.blacker@temple.edu

"two-stage" model of the AB proposes that all items in the stream are rapidly detected but then must be "consolidated" via a capacity-limited stage in order to achieve a durable, reportable state (Chun & Potter, 1995). The slow process of such consolidation leads to a perceptual "bottleneck". Consistent with the possibility that perceptual expertise might attenuate this bottleneck, faces—a category considered to be a domain of expertise for most individuals—are immune to the AB (Awh et al., 2004). Here, we first replicate the attenuation of the AB for faces and then examine whether this finding extends to other domains of expertise.

EXPERIMENT 1

Twenty-three individuals (13 males) with normal or corrected-to-normal vision participated (age, $M = 19.9$, $SD = 1.4$). Stimuli were greyscale images of 56 faces, 56 cars, 112 watches, and 112 items of furniture.

Each trial began with a fixation cross (500 ms), immediately followed by a 20 image sequence (68 ms/image). Interspersed among 18 images of furniture were two targets (Figure 1a). The first target (T1) was a watch; the second target (T2) was either a face or a car. Following each image sequence, participants indicated whether the watch had numbers or nonnumbers on its face (50% probability each). When T2 was a face, participants performed a gender discrimination task; when T2 was a car, they reported if it was a sedan or SUV. T2 randomly appeared at lags 3, 5, 9, or 11 following T1 (with stimulus onset asynchronies of 204 ms, 340 ms, 612 ms, and 748 ms, respectively). Participants performed eight blocks of 28 trials, with four alternating blocks of each T2 type.

Results

Participants with a T1 accuracy of less than 80% were excluded from all analyses ($n = 2$). Analyses of T2 performance were based only on trials in which T1 was accurately identified. Separately for face and car conditions, data were collapsed over lags 3 and 5 (inside the blink), faces, $t(20) = -1.286$, $p = .213$; cars, $t(20) = -0.26$, $p = .98$. Data were not collapsed over lags 9 and 11 due to a significant difference for the car condition, $t(20) = -2.428$, $p = .025$. Instead, lag 11 was defined as the data point outside the AB for cars and faces. A 2 (lag) × 2 (T2 category) repeated-measures ANOVA yielded a significant main effect of lag, $F(1, 20) = 22.498$, $p < .001$, with higher accuracy at longer lags, consistent with an AB. There was no main effect of T2 category, $F(1, 20) = 0.25$, $p = .623$. However, the lag × T2 category interaction was significant, $F(1, 20) = 14.117$, $p < .01$, with a greater AB for cars than for faces (Figure 1b).

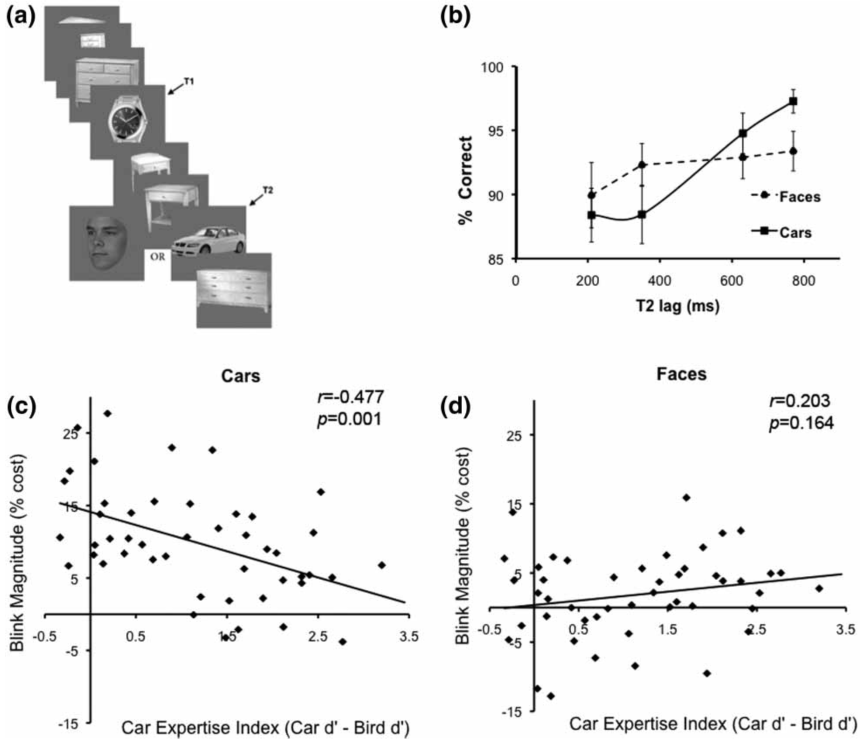


Figure 1. (a) An RSVP trial with example distractor and target stimuli. (b) T2 performance among car novices as a function of the lag from T1 for cars and faces in Experiment 1. Scatter plots showing the relationship between the car expertise index and the magnitude of the AB for cars (c) and faces (d) in Experiment 2.

Thus, consistent with previous findings, the AB was attenuated for face as compared to nonface stimuli.

EXPERIMENT 2

Fifty-five individuals (50 male) with a range of experience in identifying cars and normal or corrected-to-normal vision participated (age, $M = 21.5$, $SD = 2.5$). Perceptual expertise with cars was quantified using a previously established measure (e.g., Gauthier, Curby, Skudlarski, & Epstein, 2005), in which participants made same/different judgements about car images at the model level, regardless of year. To provide a baseline measure of perceptual skills, participants made same/different judgements about birds at the level of species. A car expertise index was defined as (car d'–bird d').

The same AB paradigm and stimuli were used as in Experiment 1, except the 28 SUVs were replaced with 28 sports cars and the T2 car judgement was changed to a discrimination task between sports cars and sedans. This change was made to avoid a potential ceiling effect on car trials among car experts.

Results

Participants with less than 80% accuracy on T1 ($n = 8$) were excluded from any analyses. Data were collapsed over lags 3 and 5 (inside the blink), cars, $t(46) = -0.258$, $p = .798$; faces, $t(46) = 0.554$, $p = .583$, and lags 9 and 11 (outside the blink), cars, $t(46) = 1.674$, $p = .101$, faces: $t(46) = -1.315$, $p = .195$, because there were no significant differences. A 2 (lag) × 2 (T2 category) repeated measures ANCOVA was used with car expertise index as the covariate of interest. Significant main effects of lag, $F(1, 45) = 65.836$, $p < .001$, and T2 category, $F(1, 45) = 120.454$, $p < .001$, emerged with performance being higher outside the attentional blink and for faces as compared to cars, respectively. The T2 category × expertise index, $F(1, 45) = 38.747$, $p < .001$, and the T2 category × lag, $F(1, 45) = 37.636$, $p < .001$, interactions were significant. The expertise index × lag interaction, $F(1, 45) = 3.815$, $p = .057$, did not reach significance. Additionally, the three-way T2 category × lag × expertise index interaction was significant, $F(1, 45) = 10.794$, $p < .01$.

To explore this three-way interaction further, a correlation analysis examined the degree to which car expertise predicted AB magnitude (i.e., [accuracy outside the blink] – [accuracy inside the blink]). Car expertise predicted the magnitude of the AB for cars, $r = -.477$, $p = .001$ (Figure 1c), but not faces, $r = .203$, $p = .164$ (Figure 1d).

DISCUSSION

Our results suggest, at least in part, that the previously documented attenuation of the AB for faces might arise as a consequence of perceptual expertise with this category, as other objects of expertise show the same benefit. Notably, degree of attenuation of the AB for car stimuli was predicted by individuals' perceptual expertise with cars. Thus, perceptual expertise appears to ameliorate for perceptual processing bottlenecks associated with awareness.

REFERENCES

Awh, E., Serences, J., Laurey, P., Dhaliwal, H., van der Jagt, T., & Dassonville, P. (2004). Evidence against a central bottleneck during the attentional blink: Multiple channels for configural and featural processing. *Cognitive Psychology*, *48*, 95–126.

Broadbent, D. E., & Broadbent, M. H. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception and Psychophysics*, *42*(2), 105–113.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in parid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109–127.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94–107.

Gauthier, I., Curby, K. M., Skudlarski, P., & Epstein, R. (2005). Individual differences in FFA activity suggest independent processing at different spatial scales. *Cognitive, Affective, and Behavioral Neuroscience*, *5*(2), 222–234.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860.

Perceptual differences between natural scene categories

Eamon Caddigan

University of Illinois at Urbana-Champaign, Champaign, IL, USA

Dirk B. Walther

Ohio State University, Columbus, OH, USA

Li Fei-Fei

Stanford University, Stanford, CA, USA

Diane M. Beck

University of Illinois at Urbana-Champaign, Champaign, IL, USA

Images of natural scenes are quickly and easily categorized by human observers (Green & Oliva, 2009; Potter & Levy, 1969; Walther, Caddigan, Fei-Fei, & Beck, 2009). Recent results have shown that images rated as good exemplars for their category are categorized more easily (Torralbo et al., 2009). Although such findings are in line with previous findings on stimulus typicality (Rosch, Simpson, & Miller, 1976), the ultrarapid speed with which observers can categorize natural scenes (VanRullen & Thorpe, 2001) raises the

---

Please address all correspondence to Eamon Caddigan, 603 East Daniel St., Champaign, IL 61820, USA. E-mail: ecaddiga@illinois.edu

possibility that good and bad exemplars of natural scenes may differ at a more perceptual stage. For instance, if scenes are classified on the basis of a feedforward sweep of information through cortex, then it is possible that those early organizational processes may be sensitive to the degree to which an image exemplifies its category, resulting in greater perceptual fluency for good exemplars of natural scenes than bad. In the following experiments, we first assess whether briefly presented good exemplars of natural scene categories are actually seen better than bad ones, and then ask whether the representation of good exemplars are more robust to degradation than bad exemplars.

EXPERIMENT 1

We first asked whether viewers "see" good scene category exemplars better than bad ones with a discrimination task that did not require categorization of the images. Participants were briefly shown a colour image of a natural scene that was either intact or 100% phase-scrambled and asked to indicate whether the image was intact or scrambled. Importantly, images were followed by a perceptual mask (500 ms duration) and image durations were staircased (23–63 ms) so that participants could correctly discriminate only 70% of the images. The images were drawn from six categories—beaches, city streets, forests, mountains, highways, and offices—and had been previously judged to be either good or bad exemplars of their categories in a separate study (Torralbo et al., 2009). After making the intact/scrambled response to each trial, participants were asked to provide a rating, on a 1–5 scale, of how clearly they saw the image.

If participants were able to see good images better than bad ones, we predicted that their clarity ratings and their performance on the intact/scrambled judgement would be higher for good than bad exemplars. Sensitivity on the intact/scrambled task, as measured by d', was significantly higher for good category exemplars than bad category exemplars (2.46 vs. 2.20), $t(19) = 3.64$, $p < .01$. A subsequent analysis showed that this was driven by miss rates for intact images (11% vs. 15%), $t(19) = 4.95$, $p < .01$, and not false alarms to scrambled images (17% vs. 18%), $t(19) = 1.24$, $p >$ .05, indicating that observers were seeing the intact images of bad category exemplars as intact less often than intact good images. Clarity rating scores, which provide a measure of participants' subjective experience of the image clarity, further confirm that participants see the good exemplars (3.80) better than the bad exemplars (3.59), $t(19) = 5.93$, $p < .01$.

EXPERIMENT 2

The fact that good exemplars were seen more clearly than bad exemplars suggests that good exemplars of a natural scene category are represented more robustly by the visual system. If this is the case then good exemplars should be more immune to visual degradation than bad exemplars. In our second experiment, we manipulated the amount of information in the scenes, and therefore their ability to conform to perceptual templates that may play a role in scene perception. We used line drawings created from the photographs used in Experiment 1, as recent work has shown that categorization of line drawings can be systematically disrupted by increasing levels of degradation (Walther, Chai, Caddigan, Beck, & Fei-Fei, 2010; see Figure 1). We controlled the amount of scene information displayed in each trial by randomly deleting line segments in the drawing: Images were displayed with 25%, 50%, or 100% of their lines.

Participants performed a categorization task on line drawings of natural scenes. On each trial, a sequence of three images was presented; the first and third “mask” images belonged to one of six categories, and the second

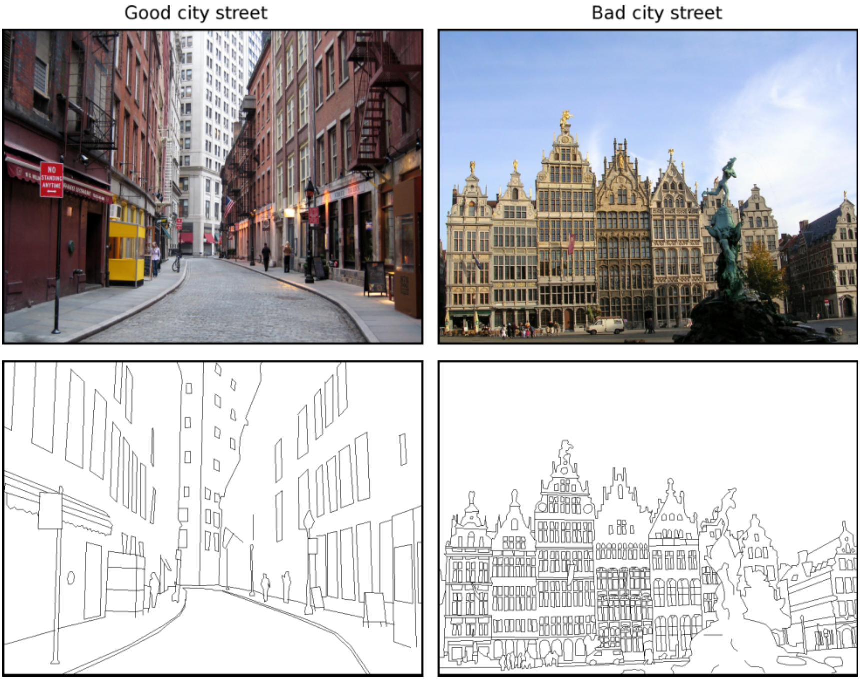


Figure 1. Photograph and line drawing of a good and bad exemplar of the “city streets” category. [To view this figure in colour, please see the online issue of the Journal.]

"target" image belonged to either the same or a different category. The masks appeared as white lines on a black background with 100% of the lines present and at a constant duration of 300 ms. The target was presented as black lines on a white background, with a variable number of lines removed, and at a brief duration (105–426 ms) determined through staircasing. Targets had previously been rated as either good or bad exemplars of their category, while masks belonged to a separate set of images with intermediate category membership ratings. Participants were asked to make a same/different category judgement about the target image in relation to the masks.

If the extent to which a scene matches a category prototype influences how well it is seen, performance should be better for good than bad category exemplars, and good images should be more immune to line deletion than bad ones. Participants' d' values for the same/different judgement were higher for the good category exemplars, $F(1, 13) = 171.11$, $p < .01$, as well as for images with a higher proportion of their lines present, $F(2, 26) = 53.69$, $p < .01$. Moreover, there was a significant interaction between these two factors, $F(2, 26) = 5.24$, $p < .05$, with good images reaching ceiling-level performance with only 50% of their lines, and bad images continuing to improve as the proportion of lines present increased.

CONCLUSION

In Experiment 1, we found that viewers "see" good natural scene category exemplars better than bad ones through an intact versus scrambled discrimination task and subjective reports of image clarity. Experiment 2 showed that line drawings of good category exemplars were easier to categorize than bad ones, and that they were more robust to degradation. Taken together, these results indicate that the extent to which an image of a scene exemplifies a category plays an important role in how well it can be apprehended by the visual system.

REFERENCES

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*(4), 464–472.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, *81*(1), 10–15.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(4), 491–502.

Torralbo, A., Chai, B., Caddigan, E., Walther, D., Beck, D. M., & Fei-Fei, L. (2009, May). *Categorization of good and bad examples of natural scene categories*. Poster presented at the ninth annual meeting of the Vision Sciences Society, Naples, FL.

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*(4), 454–461.

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience*, *29*(34), 10573–10581.

Walther, D. B., Chai, B., Caddigan, E., Beck, D. M., & Fei-Fei, L. (2010, May). *fMRI decoding of natural scene categories from line drawings.* Talk at the 10th annual meeting of the Vision Sciences Society, Naples, FL.

Human search strategies are informed by complex target distribution statistics

Matthew S. Cain

Center for Cognitive Neuroscience; Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

Ed Vul

Department of Psychology, University of California, San Diego, CA, USA

Kait Clark and Stephen R. Mitroff

Center for Cognitive Neuroscience; Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

Many important visual searches (e.g., those conducted by radiologists, baggage screeners, and military personnel) are multiple-target searches—searches where more than one target can be simultaneously present in the same search array. For example, X-ray images contain an unknown and unbounded number of potential abnormalities. Multiple-target searches are especially error prone, and may be responsible for up to half of missed abnormalities in radiology (see Berbaum, Franken, Caldwell, & Schartz, 2010, for a review). Given the uncertainty over the number of present targets, an efficient searcher must rely on assumptions about the distribution of targets across displays. Evidence suggests that multiple-target search errors are, in part, caused by such expectations about the likelihood of a target being present, with the less likely targets being missed more often (e.g., Fleck, Samei, & Mitroff, 2010). Similarly, in searches for multiple categories of targets (where only one target ever appears within a given array), targets from the rarer category are more likely to be missed (e.g., Menneer,

---

Please address all correspondence to Matthew S. Cain, B203 Levine Science Research Center, Box 90999, Duke University, Durham, NC 27708, USA. E-mail: matthew.s.cain@duke.edu

Donnelly, Godwin, & Cave, 2010), and target prevalence influences a searcher's decision criterion, leading to more false alarms with high prevalence and more misses with low prevalence (Wolfe & van Wert, 2010).

Although early termination of search can cause errors, it is not necessarily suboptimal: Optimal searchers *should* adjust their strategy according to the distribution of targets across displays and environmental pressures. A large body of evidence suggests that people are aware of higher order environmental statistics and can adapt their behaviour accordingly. For example, the dynamics of memory practice effects, forgetting rates, and training spacing effects appear well matched to the frequencies with which such information is needed (Anderson & Schooler, 1991).

Here we explore whether searchers are sensitive to complex target distribution statistics and whether they adapt their strategies accordingly. In particular, if people learn that targets tend to co-occur—most displays contain few or zero targets, whereas rare displays contain many targets—do they use this information to decide how long to search a given display? Three separate groups of searchers saw different distributions of the number of targets present in each display. These distributions all had the same target prevalence, but differed in whether the targets tended to co-occur in a few displays (randomly presented across the experiment), or whether targets were dispersed over many displays. We predicted that each group's performance would reflect the statistics of their search environment such that search termination times would reflect both the learned tendency of targets to cluster in displays and the evidence acquired during searching about whether a particular display contains a target cluster.

METHODS

Thirty-three students from Duke University were divided randomly into three groups of 11. Each trial presented a search array of 40 items on a cloudy background. Targets were perfect "T" shapes and the remaining items were distractor "L" shapes that were not perfect Ts. Each trial had 0–12 targets, with target prevalence manipulated between groups (see Figure 1A). Participants clicked on each T they found then clicked a button labelled "Done", ending the trial. Feedback was given after each trial to provide participants the search environment statistics, regardless of their performance. Participants received 15 points for each target found and the experiment ended at 2000 points.

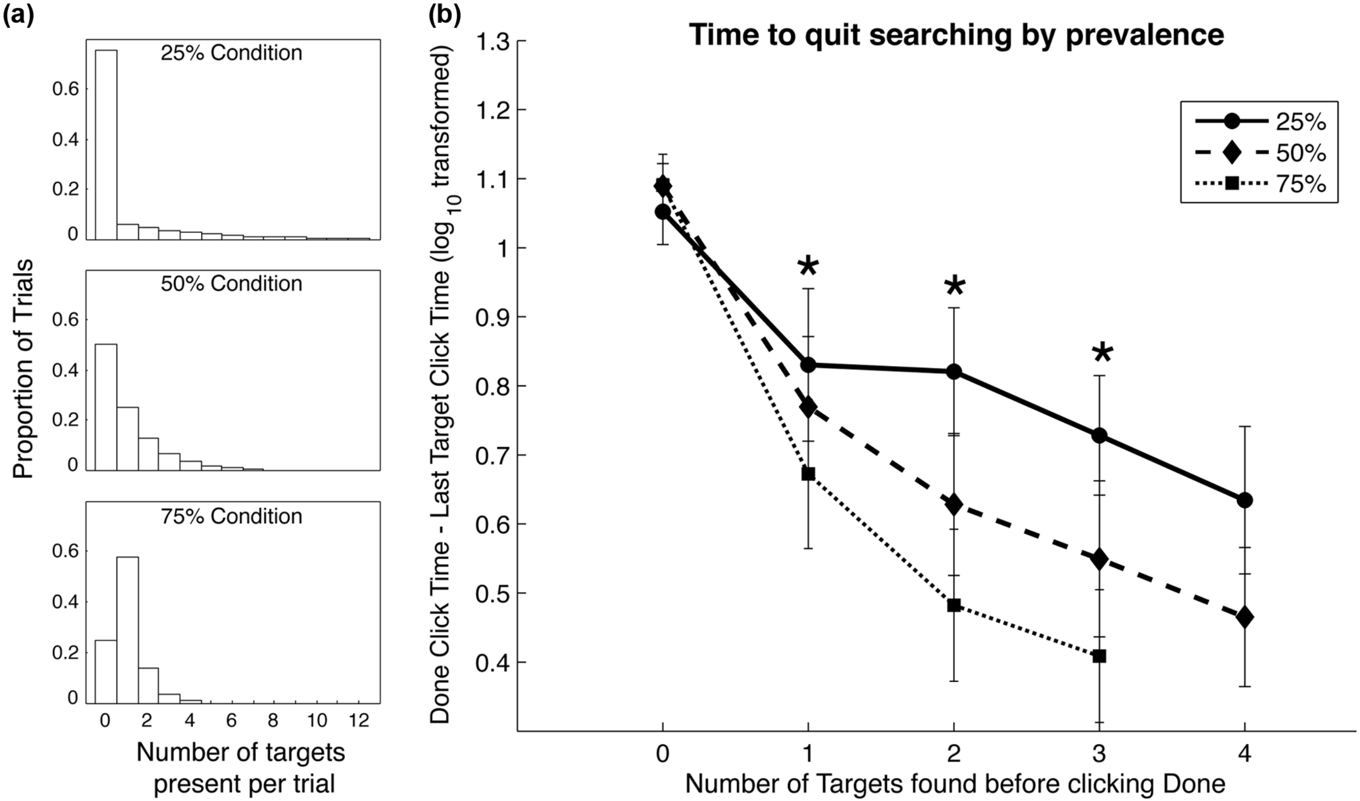


Figure 1. (A) Proportion of trials with each number of targets by condition. (B) Time spent searching after clicking the last target found, plotted against the number of targets found. Not enough participants contributed data for four targets found in the 75% condition to be informative because those trials were rare by design.

Conditions

Target prevalence was fixed across conditions, so participants saw the same average number of targets per display (one) in all conditions. However, we varied whether targets tended to cluster across conditions. Displays had either a 25%, 50%, or 75% chance of containing a target. The number of targets present in each trial was sampled from a geometric distribution, providing a search environment with complex but informative target prevalence statistics. Thus, in the 25% condition (compared to the 50% and 75% conditions) few displays contained targets, but those displays were likely to contain *more* targets. Conversely, in the 75% condition relatively more displays contained targets, but those displays were likely to contain relatively *few* targets. Thus, our effects are driven by how targets were distributed across trials, not by overall target prevalence.

RESULTS

Our primary measure is time spent searching after a target has been found—the time between the last target actually found on a trial (even if not all targets were found) and when the "Done" button was clicked. Figure 1B plots this difference measure for each condition against the number of targets found. The pattern is clear: The more targets that were *likely* to be in a display, the longer participants continued searching.

We conducted a 3 × 3 repeated measures ANOVA on the log-transformed RT differences with condition (25%, 50%, or 75%) and number of targets found (0, 1, or 2) as factors (not all participants found three or more targets). There were main effects of number of targets found, $F(2, 30) = 203.65$, $p < .001$, and condition, $F(2, 30) = 6.76$, $p = .004$. The key statistical comparison was the interaction between condition and number of targets found, $F(4, 30) = 12.22$, $p < .001$. With clustered targets (25% condition), participants were slow to end their search once they had found at least one target. Conversely, with dispersed targets (75% condition), participants were quick to end their search after finding a target. Performance in the 50% condition was intermediate, suggesting sensitivity to target distributions is continuous, rather than engaging discrete "rare target" and "frequent target" modes.

DISCUSSION

The present experiment demonstrates that searchers are sensitive to the statistics of their search environment: Participants terminated their searches quickly when finding an additional target was unlikely but searched longer when finding an additional target was more likely. This was observed both within participants (with a main effect of number of targets found), and between groups (with an interaction between condition and number of targets found), suggesting that people adapt and optimize their search strategies to match the complex statistics of the environment. This has broad implications for search, suggesting that artificially modifying target distribution statistics, such as priming baggage screeners with daily training runs of multiple-target bags (cf. Wolfe et al., 2007), may be an effective way to enhance sensitivity in critical multiple-target visual searches.

REFERENCES

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*(6), 396–408.

Berbaum, K. S., Franken, E. A., Jr., Caldwell, R. T., & Schartz, K. M. (2010). Satisfaction of search in traditional radiographic imaging. In E. Samei & E. Krupinski (Eds.), *The handbook of medical image perception and techniques* (pp. 107–138). Cambridge, UK: Cambridge University Press.

Fleck, M. S., Samei, E., & Mitroff, S. R. (2010). Generalized "satisfaction of search": Adverse influences on dual-target search accuracy. *Journal of Experimental Psychology: Applied*, *16*(1), 60–71. doi: 10.1037/a0018629

Menneer, T., Donnelly, N., Godwin, H. J., & Cave, K. R. (2010). High or low target prevalence increases the dual-target cost in visual search. *Journal of Experimental Psychology: Applied*, *16*(2), 133–144. doi: 10.1037/a0019569

Wolfe, J. M., Horowitz, T. S., van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General*, *136*(4), 623–638. doi: 10.1037/0096-3445.136.4.623

Wolfe, J. M., & van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology*, *20*(2), 121–124. doi: 10.1016/j.cub.2009.11.066

Perceptual grouping determines the locus of attentional selection

Joshua D. Cosman and Shaun P. Vecera

Departments of Neuroscience and Psychology, University of Iowa, Iowa City, IA, USA

Selective attention allows us to process task-relevant information while effectively ignoring task-irrelevant information. For example, our ability to

---

Please address all correspondence to Joshua Cosman, Departments of Neuroscience and Psychology, E305 Seashore Hall, University of Iowa, Iowa City, IA 52242-1407, USA. E-mail: joshua-cosman@uiowa.edu

read a newspaper in a noisy, crowded coffeehouse depends on our ability focus on the words on the page while simultaneously ignoring the conversations and sounds around us. Although much is known about the effects of attentional selection, the locus of processing at which such selection occurs (i.e., early vs. late in processing) is long debated (Duncan, 1980; Lachter, Forster, & Ruthruff, 2004; Treisman, 1969). As a resolution, Lavie and colleagues have proposed that the locus of attentional selection is flexible, being determined by the demands, or *perceptual load*, of task-relevant information processing (Lavie, 1995; Lavie, Hirst, de Fockert, & Viding, 2004). Specifically, perceptual level attention is viewed as a finite resource—when perceptual load is high and processing capacity is exhausted, early selection is induced and the processing of task-irrelevant distractors is attenuated at an early level of processing. Conversely, when perceptual load is low, there are sufficient attentional resources left to "spill over" and process task-irrelevant distractors. Given its parsimonious resolution to the debate regarding the locus of selection, load theory has been an influential theory of attentional selection in both cognitive psychology and neuroscience, being supported by numerous behavioural and neurophysiological studies.

In the current experiment, we tested whether the locus of selection, as measured by perceptual load effects, can be modulated by perceptual grouping. Given that perceptual grouping serves up objects that control the allocation and spread of attentional resources (e.g., Richard, Lee, & Vecera, 2008; Vecera & Farah, 1994), it is plausible that perceptual grouping might directly influence the level at which selective attention exerts its effects. For example, features of task-relevant objects may be obligatorily processed under high-load conditions even when the features themselves are task irrelevant, and features of task-irrelevant objects may be effectively ignored under low-load conditions. In support of this possibility, there is evidence that perceptual grouping can modulate the processing of task-irrelevant information under some conditions (e.g., Baylis & Driver, 1992; see also Chen, 2003).

We addressed this possibility by having observers perform a search task in which we varied both perceptual load and the position of task-irrelevant distractors relative to the search arrays—specifically, the task-irrelevant flanker letter could appear either in the same object as the search array or in a different object. With this design it was possible to examine the effect of perceptual grouping on processing under different conditions of load. If perceptual grouping modulates the locus of attentional selection, we would expect to see interference emerge when the flanker is contained within the same object as the search array, but not when it appears in a different object than the search array, regardless of perceptual load. In other words, it is possible that perceptual grouping, not perceptual load, may determine of the locus of attentional selection.

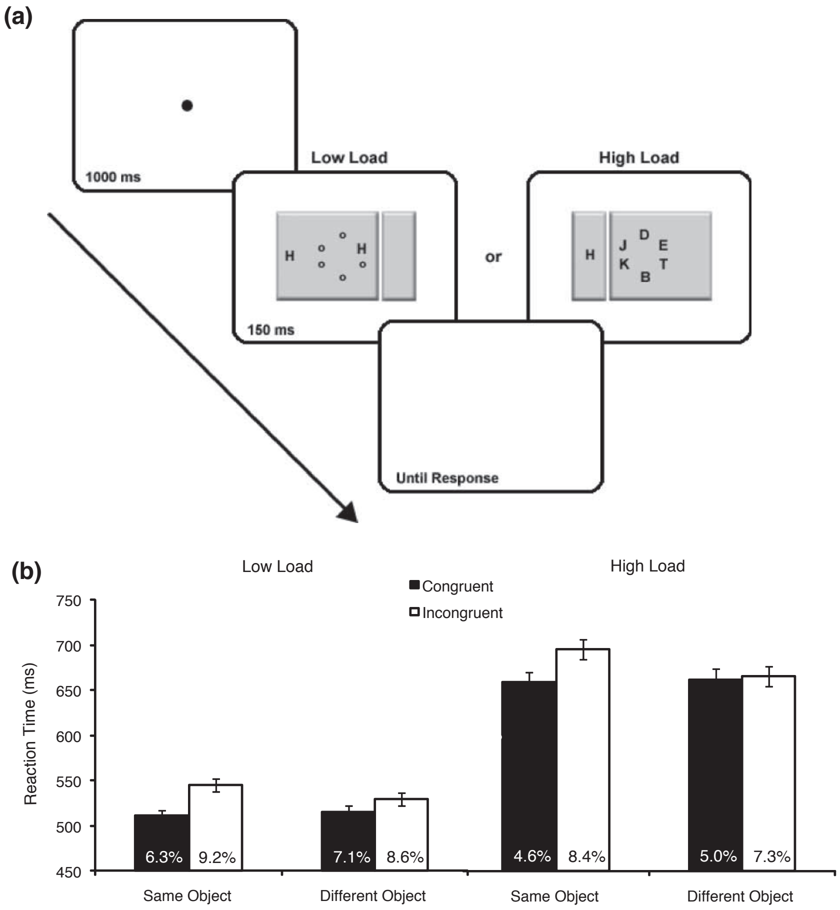


Figure 1. (a) The trial sequence, giving example of low-load different object (left) and high-load same object trials (right). (b) Reaction times and error rates for each condition in the experiment. Error bars represent 95% confidence intervals (Loftus & Masson 1994).

METHOD

Sixteen University of Iowa undergraduates performed a basic search task (Figure 1). Following the presentation of a fixation point for 1000 ms, a search display was presented for 100 ms. The displays consisted of two grey 3-D rendered objects on a white background, one large (12° × 10°) and one small (4° × 10°). The large object always contained the task-relevant search array, and on half of the trials also contained a single, task-irrelevant flanker letter (same-object flanker condition). On the other half of trials, the flanker letter appeared in the smaller object (different-object flanker condition). The relative location of each object (left vs. right side of display) was randomly determined on a given trial, with the task-irrelevant flanker letter being either congruent or incongruent on a given trial. The search arrays were either high-load displays containing a target letter (E or H) among five heterogeneous distractor letters (D, J, K, B, and T, each measuring 0.9° × 1.4°), or low-load displays consisting of the target letter and five small placeholder circles (0.5° diameter), with load being blocked (cf. Lavie & Cox, 1997). Participants were told to maintain central fixation, and search the circular arrays for the target while ignoring the task-irrelevant flankers and objects. Participants performed three high- and three low-load blocks of 96 trials each, with load blocks alternated and order counterbalanced across subjects.

RESULTS

An omnibus ANOVA with flanker object (same vs. different) display load (high vs. low), and flanker congruency (congruent vs. incongruent) was performed on correct RTs. We observed main effects of congruency, $F(1, 15) = 33.5$, $p < .0001$, and load, $F(1, 15) = 83.6$, $p < .0001$, as well as a significant interaction between flanker object and congruency, $F(1, 15) = 11.0$, $p < .01$. No other main effects or interactions were significant, $F\text{s} < 3.5$, $p\text{s} > .08$. Secondary two-way ANOVAs were conducted on RTs from high and low load conditions individually to examine the root of the flanker object by congruency interaction. Importantly, significant two-way interactions between flanker object and congruency were observed in both the high-load, $F(1, 15) = 7.7$, $p = .01$, and low-load, $F(1, 15) = 4.9$, $p = .04$, conditions, indicating that flanker effects were significantly *larger* when the flanker appeared in the same object as the target, regardless of load.

DISCUSSION

Our results show for the first time that perceptual grouping is a major determinant of the locus of attentional selection, flexibly increasing or decreasing filtering efficiency based on whether the task-relevant and irrelevant information are part of the same perceptual group. During high-load search, where attentional capacity should be exhausted and attentional filtering very effective (Lavie, 1995; Lavie et al., 2004), task-irrelevant flanker letters still exert an interference effect so long as they group with the task-relevant search array. Conversely, during low-load search, filtering efficiency is increased when the to-be-ignored letter does not group with the search array. Given this direct modulation of perceptual load effects by grouping, it appears that perceptual grouping, rather than perceptual load, is the primary determinant of what information is processed and allowed to affect behaviour.

REFERENCES

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception and Psychophysics*, *51*, 145–162.

Chen, Z. (2003). Attentional focus, processing load, and Stroop interference. *Perception Psychophysics, 65*, 888–900.

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

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty years after Broadbent: Still no identification without attention. *Psychological Review*, *111*, 880–913.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 451–468.

Lavie, N., & Cox, S. (1997). On the efficiency of attentional selection: Efficient visual search results in inefficient rejection of distraction. *Psychological Science*, *8*, 395–398.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339–354.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review*, *1*, 476–490.

Richard, A., Lee, H., & Vecera, S. P. (2008). Attentional spreading in object-based attention. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 842–853.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*, 282–299.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, *123*, 146–160.

Unitary vs. multiple attentional loci reflect space-based vs. object-based modes of attention

Lisa N. Jefferies and Steven Yantis
Johns Hopkins University, Baltimore, MD

Visual stimuli are processed faster and more accurately when they appear at attended locations. Theories of spatial attention tend to appeal either (1) to the idea of a single unitary focus of attention that expands and contracts to optimize performance on the task at hand (e.g., Eriksen & Yeh, 1985), or (b) to multiple foci deployed to different locations simultaneously (e.g., Awh &

Please address all correspondence to Lisa N. Jefferies, Department of Psychological and Brain Sciences, Johns Hopkins University, 136 Ames Hall, 3400 North Charles Street, Baltimore, MD 21218, USA. E-mail: ljefferi@gmail.com

Pashler, 2000). Recently, Jefferies, Enns, and Di Lollo (2010) have argued that observers can flexibly deploy either mode of attention depending on mental set and can switch between modes during a task. To date, research investigating whether the focus of attention is unitary or divided has assumed the two modes to be simply two alternative distributions of attention across space. Here we ask whether this assumption is justified or whether, in fact, unitary and divided attention are fundamentally different in nature. Specifically, we probe whether a unitary focus of attention reflects space-based attention, whereas a divided focus of attention reflects object-based attention.

We based our methodology on that of Jefferies et al. (2010). In that research, attention was indexed by means of two well-established measures of attention: the attentional blink and Lag-1 sparing. The attentional blink refers to the impaired identification of the second of two rapidly sequential targets (T1, T2; Raymond, Shapiro, & Arnell, 1992); Lag-1 sparing refers to the paradoxical finding that identification of T2 is unimpaired when it appears directly after T1 (Potter, Chun, Banks, & Muckenhoupt, 1998). Since Lag-1 sparing only occurs if T2 appears within the focus of attention (Jefferies, Ghorashi, Kawahara, & Di Lollo, 2007), its occurrence can be used to probe whether or not T2 appears within an attended region and hence whether the focus of attention is unitary or divided.

In Jefferies et al. (2010), observers were presented with two concurrent RSVP streams of digit distractors, one on either side of fixation, separated by a blank region. Two pairs of letter targets (T1-pair, T2-pair) could appear either within the RSVP streams or in the central blank region between the streams. By varying whether the location of the T1-pair was predictable (always in-stream) or not (unpredictably in-stream or between-streams), observers were encouraged to deploy either a divided or a unitary focus, respectively. In both conditions, the T2-pair subsequently appeared unpredictably either within the streams or in the central region. In those critical trials in which the T2-pair appeared between-streams, Lag-1 sparing occurred in the unpredictable condition (indicating an attended central region, consistent with a unitary focus of attention) but not in the predictable condition (indicating an unattended central region, consistent with a divided focus of attention).

The goal of the current study was to determine whether the deployment of attention as a unitary or a divided focus simply reflects two different patterns of deployment across space, or whether unitary and divided attention in fact reflect space- and object-based attention, respectively. To test this hypothesis, we added a secondary, working memory task to the procedure of Jefferies et al. (2010). That is, prior to the onset of the RSVP streams, observers were shown an array of items and required to maintain either the spatial location or the colours of those items, thus selectively

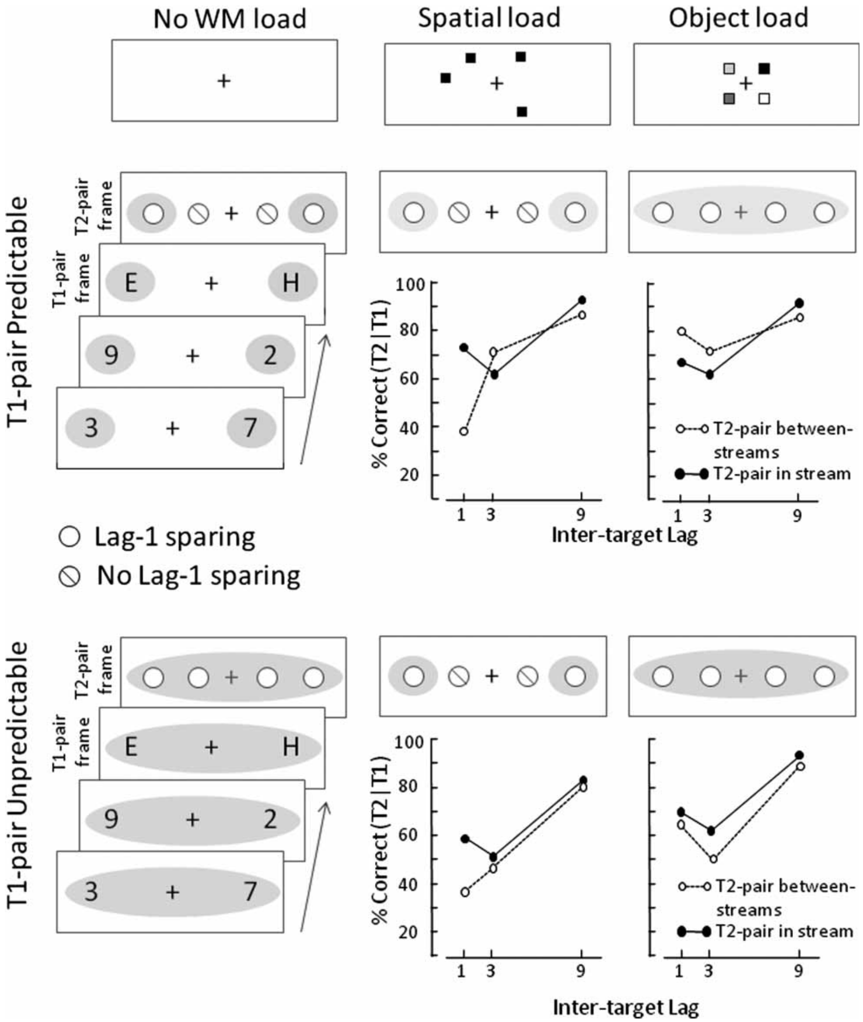


Figure 1. Schematic representation of the sequence of events within a trial. Observers were first presented with the working memory array, followed by the attentional blink task. The hypothesized distribution of the focus of attention is represented by the pale grey ovals and is shown to be either unitary or divided depending on whether the T1-pair is predictable or unpredictable and on the nature of the working memory load. The greyscale patches in the object working memory task represent squares of different colours.

taxing either the observers' spatial processing resources or their object processing resources (see Figure 1; cf. Oh & Kim, 2004). If the deployment of a unitary focus of attention requires spatial processing resources, then a difficult concurrent task that utilizes those same spatial resources should impair the deployment of a unitary focus of attention. In this case, evidence should be found for a divided focus. Similarly, if the deployment of a divided focus of attention requires object processing resources, then a concurrent task that also utilizes object processing resources should force observers to instead deploy a unitary focus.

The two variables, T1-pair predictability and type of working memory load, were crossed in a between-subjects design. In determining whether the focus of attention is unitary or divided, the trials of interest are those in which the T2-pair appears between-streams. Lag-1 sparing is present if T2-pair accuracy is higher at Lag 1 than Lag 3; the opposite pattern indicates a lack of Lag-1 sparing. The results unequivocally supported our hypotheses. In the T1-pair predictable condition (divided focus most efficient), no Lag-1 sparing occurred when the T2-pair appeared between-streams if observers were given a spatial working memory load, $t(10) = -3.8$, $p < .01$. The absence of Lag-1 sparing indicates an unattended central region, consistent with a divided focus of attention. In other words, a spatial working memory load did not interfere with the deployment of a divided focus of attention. In contrast, when observers were given an *object* working memory load, strong Lag-1 sparing was found, indicative of an attended central region and a unitary focus of attention, $t(14) = 2.73$, $p < .02$. In this case, the object working memory load prevented the deployment of a divided focus of attention, forcing observers to instead deploy a unitary focus of attention. In the T1-pair unpredictable condition (unitary focus most efficient), Lag-1 sparing was found when the T2-pair appeared between-streams if the concurrent working memory load contained object information, $t(9) = 2.49$, $p < .05$, but was strikingly absent if the load contained spatial information, $t(11) = -3.61$, $p < .01$. In this case, observers were unable to maintain spatial information and simultaneously deploy a unitary focus of attention, presumably because the demands of the two tasks overlapped. Type of working memory load had no effect on T1 accuracy ($p = .48$), supporting the conclusion that the resource drain incurred by the working memory load had its primary influence on the mode of attentional deployment and not on target processing and identification.

In summary, these results suggest that different mechanisms underlie the deployment of unitary and divided foci of attention. A unitary focus seems to rely on spatial processing resources whereas a divided focus of attention requires object processing resources. Given this underlying difference, these two modes of focused attention may enhance different kinds of information processing and will be optimally deployed in different situations depending on the nature of processing resources available.

REFERENCES

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 834–846.

Eriksen, C. W., & Yeh, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 583–597.

Jefferies, L. N., Enns, J. T., & Di Lollo, V. (2010). Whether attention is unitary or divided depends on task demands. *Manuscript submitted for publication*.

Jefferies, L. N., Ghorashi, S., Kawahara, J.-I., & Di Lollo, V. (2007). Ignorance is bliss: The role of observer expectation in dynamic spatial tuning of the attentional focus. *Perception and Psychophysics*, *69*, 1162–1174.

Oh, S.-H., & Kim, M.-S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin and Review*, *11*, 275–281.

Potter, M. C., Chun, M. M., Banks, B. S., & Muckenhoupt, M. (1998). Two attentional deficits in serial target search: The visual attentional blink and an amodal task-switch deficit. Journal of Experimental Psychology: Learning, Memory. *and Cognition*, *24*, 979–992.

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

Launching curved apparent motion: A motion interpolation study

Sung-Ho Kim, Jacob Feldman, and Manish Singh

Department of Psychology, Center for Cognitive Science, Rutgers University, Piscataway, NJ, USA

One of the fundamental challenges of vision is how to extract invariant information from the changing visual stimulus caused by both object and observer motion (Gibson, 1979). Objects come into and out of view over time, but we perceive entire objects from fragments of objects occluded by others. This phenomenon has been termed “amodal completion” (Michotte, Thinès, & Crabbé, 1964/1991). Another challenge of vision is how representations of observed movements and their causal consequences are formed. When a moving object stops abruptly next to another object, and then the second object starts to move in the same direction as the first object, people can directly perceive not only object motion, but also causality, such that the first object appears to impart momentum to the second, called the launching effect (Michotte, 1946/1963).

Please address all correspondence to Sung-Ho Kim, Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854-8020, USA. E-mail: sungho4@eden.rutgers.edu

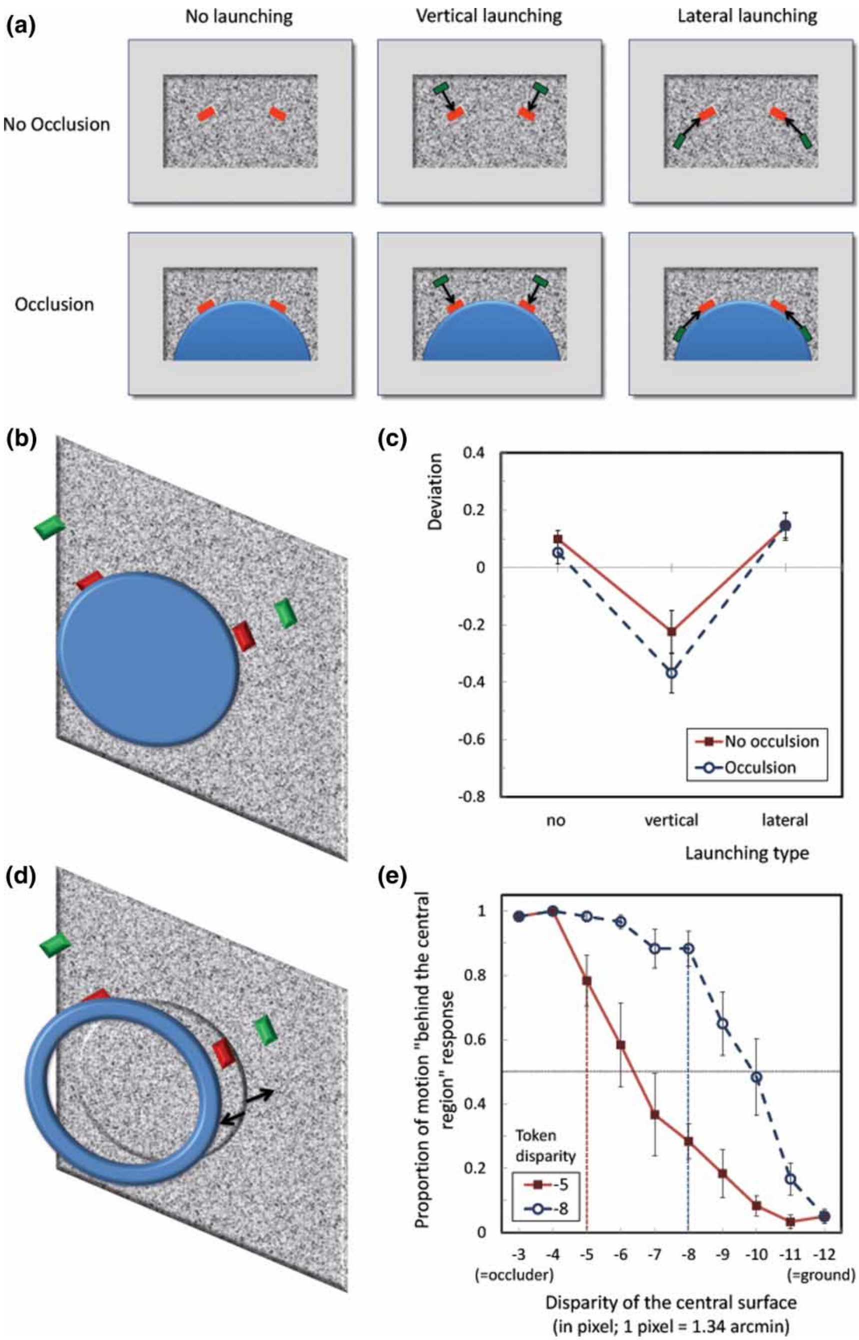


Figure 1. (Caption on next page.)

Here we report a study examining the interaction of amodal completion and causal perception with apparent motion. We hypothesized that under abrupt disappearance of an object, people would perceive it as hidden behind an adjacent occluder, and that this impression would be more likely given the momentum transferred by the launching effect. We tested this idea in an "apparent motion" task by exploiting a well-known phenomenon, the tunnel effect (Burke, 1952): The perception of a continuous movement of a single object with its path partially hidden by a screen (tunnel), despite the presence of two distinct objects.

In recent studies (Kim, Singh, & Feldman, 2010), we examined whether amodal completion and the launching effect can bias an apparent motion path towards longer curved paths behind an occluder, violating the apparently well-established principle that apparent motion follows the shortest possible path. We presented observers motion sequences containing a small token alternately abutting each end of a semicircular occluder, and asked them to determine whether they perceived a straight or a curved path motion. The longer the interstimulus interval (ISI), the more observers tended to report the curved path over the straight path. Furthermore, when two more objects were added so as to appear to collide with the motion tokens at offset, in the direction of either the straight or the curved path, the motion was almost exclusively perceived in the direction of the launch.

In the present study, we further examined the dynamic nature of amodal representation. In Experiment 1, we quantitatively measured the perceived geometry of the curved path induced by dynamic occlusion and the launching effect. In our previous study, once observers experience token motion behind the semicircular tube, its path was likely to be guided by the shape of the tube. To prevent path-guided motion, Experiment 1 employed a disk-shaped occluder. Eighteen observers viewed motion sequences of two alternating red rectangular tokens with varying ISI (100 or 400 ms) in stereoscopic displays. They indicated the perceived trajectory at the vertical midline of the perceived motion path, by adjusting the vertical location of a probe dot to the point of the perceived mid-point of the token trajectory. Their adjusted location was expressed in terms of "deviation" from the line connecting the tokens, normalized by half the intertoken distance. In an occlusion condition, a large disk was placed in front of the tokens in depth (tokens were positioned along the circumference of the disk such that they were perceived as partially occluded by the disk; see Figure 1a and b). In a no-occlusion condition, tokens were positioned the same as in the occlusion

Figure 1. (a) Launching types and occlusion conditions in Experiment 1. Arrows represent launching directions. (b) 3-D layout of stimuli used in Experiment 1. (c) Results of Experiment 1. Error bars represent ±1 SE. (d) 3-D layout of stimuli used in Experiment 2. (e) Results of Experiment 2. [To view this figure in colour, please see the online issue of the Journal.]

condition, but the disk was not present (Figure 1a). To elicit a Michotte-style "launch", we added two green rectangles, which appeared to collide with the motion tokens at offset, in a direction along the circumference of the disk (lateral launching), or orthogonal to the disk (vertical launching).

Results showed that when only the tokens were presented without launching, rotational motion was perceived, its trajectory slightly curved above the straight line connecting them (see Figure 1c). In the lateral launching condition the interpolated path was more curved towards the circumference of the disk, which is the ideal path predicted by kinematic geometry (Shepard, 1984), but in the vertical launching condition the midpoint of path was below the straight line towards the centre of the disk, contrary to kinematic geometry, $F(2, 32) = 35.16$, $p < .001$. Also these trends were stronger in the occlusion than in the no-occlusion condition, $F(1, 16) = 3.15$, $p = .095$. The effect of ISI was not significant, but the three-way interaction among launching type, occlusion, and ISI was significant, $F(2, 32) = 4.42$, $p = .02$. When ISI was long, the interpolated path in vertical launching became more curved with the presence of an occluder.

In Experiment 2, we investigated whether apparent motion behind an occluder would be experienced even when the motion path is not completely occluded. We employed a doughnut-shaped occluder, instead of a solid disk, and manipulated the binocular disparity of the texture within the central circular region (varying from the background to the occluder disparity; see Figure 1d), and the disparity of the two tokens. In a yes/no task, six observers (including author SK) indicated whether they saw the motion tokens passing behind the central circular region. The results showed that when the central circular region became close in depth to the background, observers were less likely to perceive the motion occluded by the circular area (Figure 1e). But, given sufficient depth difference between the central circular region and the background, the momentum transferred by the collision could perceptually drive the tokens behind the circular region, even though tokens were slightly (1 or 2 pixel disparity) in front of it (so that occlusion of the whole path was logically impossible).

In sum, our results suggest that (1) the amodal representation of fully hidden objects and the momentum induced by the launching effect strongly modulate the trajectory of object motion, confirming and extending our previous findings, and (2) they become integral parts of the whole motion experience such that the motion behind an occluder is experienced even when intersurface relationships do not support it.

REFERENCES

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, *4*, 121–138.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Kim, S., Singh, M., & Feldman, J. (2010, May). *Curved apparent motion induced by amodal completion and the launching effect*. Poster presented at the 10th annual meeting of the Vision Sciences Society, Naples, FL.

Michotte, A. (1963). *The perception of causality.* New York, NY: Basic Books. (Original work published 1946)

Michotte, A., Thinès, G., & Crabbé, G. (1991). Amodal completion of perceptual structures. In G. Thines, A. Costall, & G. Butterworth (Eds.), *Michotte's experimental phenomenology of perception* (pp. 140–167). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. (Original work published 1964).

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imaging, thinking, and dreaming. *Psychological Review, 194,* 417–447.

Temporal oscillations in attention capture by moving stimuli as revealed by fMRI

Jennifer R. Lechak and Andrew B. Leber

Department of Psychology, University of New Hampshire, Durham, NH, USA

In order to efficiently process our visual environment, we must segment it according to our current behavioural goals, processing relevant stimuli while discarding irrelevant stimuli. To this end, attention can act in a goal-driven way, such that we orient towards stimuli that match a known feature (e.g., colour). Also, attention can act in a stimulus-driven way, in which our attention is captured by the most salient stimuli. How the components of goal-driven and stimulus-driven control interact has long been a critical focus of research. Investigators tend to characterize attention as being dominated by one component or the other, but it is also possible that attention may operate on a spectrum between these components in such a way that sometimes processing is largely goal driven, and sometimes it is largely stimulus driven. That is, observers might succeed at avoiding distraction by salient, irrelevant stimuli during some moments; at others the observers are susceptible to robust attention capture. Previous research may have missed the possibility that attentional control fluctuates because most studies average task performance across sessions to determine the overall amount of attention capture by a distracting stimulus.

Please address all correspondence to Jennifer R. Lechak, Department of Psychology, University of New Hampshire, Durham, NH 03824, USA. E-mail: Jennifer.Lechak@unh.edu

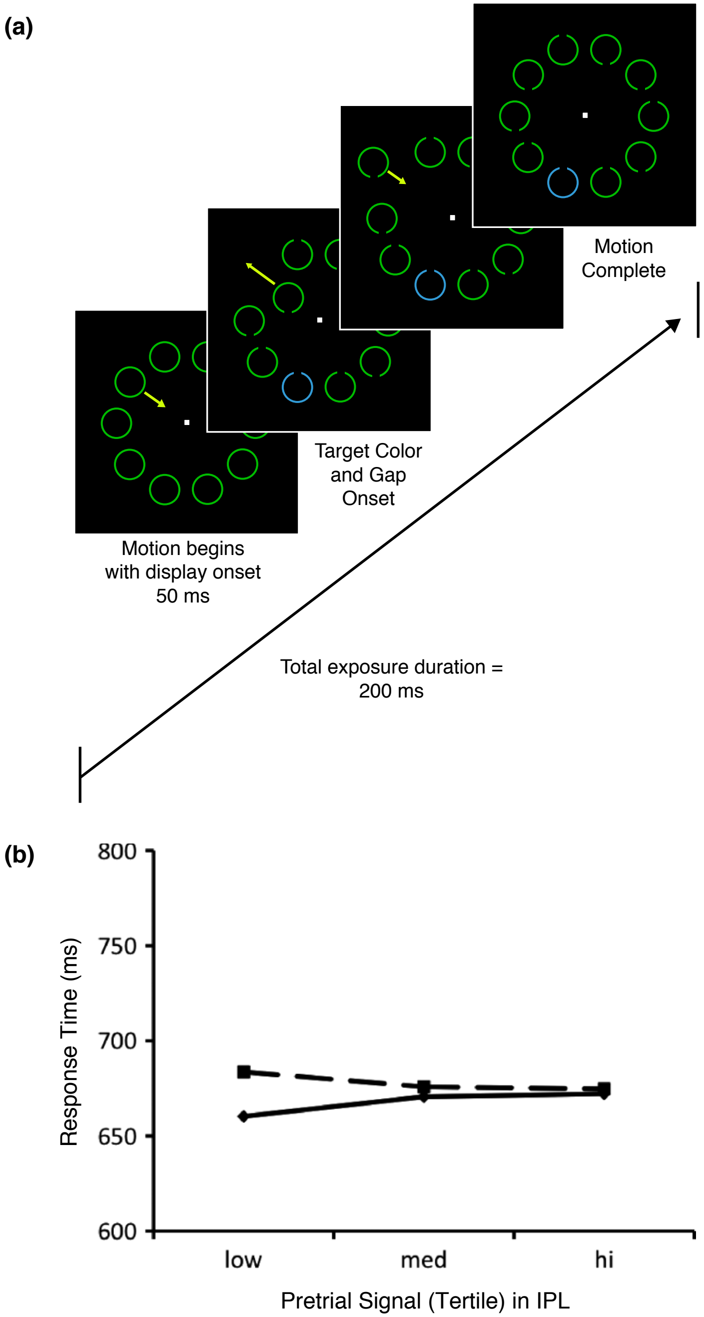


Figure 1. (Caption on next page.)

Recent work from our lab has indeed found such fluctuations in the magnitude of attention capture within an individual's performance over a period of trials (Leber, 2010). Utilizing a unique fMRI analysis technique in which temporal variations in neural activity are used to illuminate these behavioural fluctuations, behaviour can be predicted based on neural activity before the trial begins (Leber, Turk-Browne, & Chun, 2008). Leber (2010) found that distraction by colour could be predicted by pretrial fMRI signal measured from specific brain regions, namely the middle frontal gyrus (MFG). That is, fluctuations could be predicted in advance of the trials, such that higher pretrial signal was linked to reduced capture on the upcoming trial (Leber, 2010).

The current fMRI study aims to generalize the findings of Leber (2010) by studying additional stimulus features. Although it is known that fluctuations in the magnitude of capture by colour, a static stimulus feature, can be predicted by examining pretrial fMRI signal, it is unknown whether and how other forms of capture can be predicted. Here, we investigated whether pretrial signal can be used to predict behavioural distraction by a dynamic stimulus (i.e., a motion singleton). Irrelevant visual motion has previously been shown to produce robust attention capture (Abrams & Christ, 2003, Franconeri & Simons, 2003). We investigated whether fluctuations in capture by motion could be predicted, and, if so, whether the same brain regions mediated such fluctuations.

METHODS

We scanned 10 healthy participants using a Siemens TimTrio 3-Tesla MRI scanner. Participants were asked to perform a feature-based search for a colour singleton and were asked to identify the location of a gap either in the top or the bottom of the object (see Figure 1a). Participants' responses were recorded using a fibre-optic button box, the right index finger indicating a gap on the top and right middle finger indicating a gap on the bottom. A motion singleton distractor appeared on 50% of trials.

RESULTS

Results showed that moment-to-moment fluctuations in behavioural distraction could be predicted using pretrial signal. One significant cluster of voxels was observed in the inferior-parietal lobule (IPL) located at Talaraich

Figure 1. (a) Depiction of the stimulus events during the trial. (b) Mean response times plotted as a function of pretrial signal tertile in IPL, for distractor present (dashed line) and distractor absent (solid line) conditions. [To view this figure in colour, please see the online issue of the Journal.]

coordinates [−34, −35, 45]. IPL data were sorted into tertiles of low, medium, and high pretrial signal to examine the consequent behaviour in these conditions; in particular, we wished to compare distraction effects in the outermost tertiles (i.e., low vs. high). To this end, a 2 (tertile) × 2 (distractor presence) ANOVA was carried out on the RT data. A significant main effect of distractor presence was observed, $F(1, 9) = 6.739$, $p = .029$. This effect interacted significantly between low and high tertiles of pretrial signal, $F(1, 9) = 7.318$, $p = .024$, and distraction was reduced from 23 ms at the low tertile to 2 ms at the high tertile (see Figure 1b). This result shows that trial-by-trial fluctuations in pretrial activity in IPL are predictive of the current degree of attentional control within individuals during the course of an experimental session. MFG was also examined, but it did not reliably predict behavioural distraction.

DISCUSSION

The interaction between goal-driven and stimulus-driven control of attention governs how we perceive our environment. Recent work has begun to show that the nature of this interaction can fluctuate on a moment-to-moment basis (Leber, 2010), but the generality of this phenomenon has yet to be sufficiently explored. This study investigated whether pretrial fluctuations in fMRI signal could be predictive of distraction by a moving stimulus. These results replicate Leber (2010) in that they observed vacillations between goal-driven and stimulus-driven attentional control. However, such vacillations were predictable only from IPL, not from MFG. It is possible that static and dynamic stimuli utilize different brain regions for attentional control, such that MFG mediates resistance to distraction only from static stimuli, whereas IPL mediates resistance to distraction from dynamic stimuli.

REFERENCES

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science*, *14*(5), 427–432.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Attention, Perception, and Psychophysics*, *65*, 999–1010.

Leber, A. B. (2010). Neural predictors of within-subject fluctuations in attentional control. *Journal of Neuroscience*, *30*, 11458–11465.

Leber, A. B., Turk-Browne, N. B., & Chun, M. M. (2008). Neural predictors of moment-to-moment fluctuations in cognitive flexibility. *Proceedings of the National Academy of Sciences, USA*, *105*(36), 13592–13597.

Temporal dynamics of the allocation of spatial attention

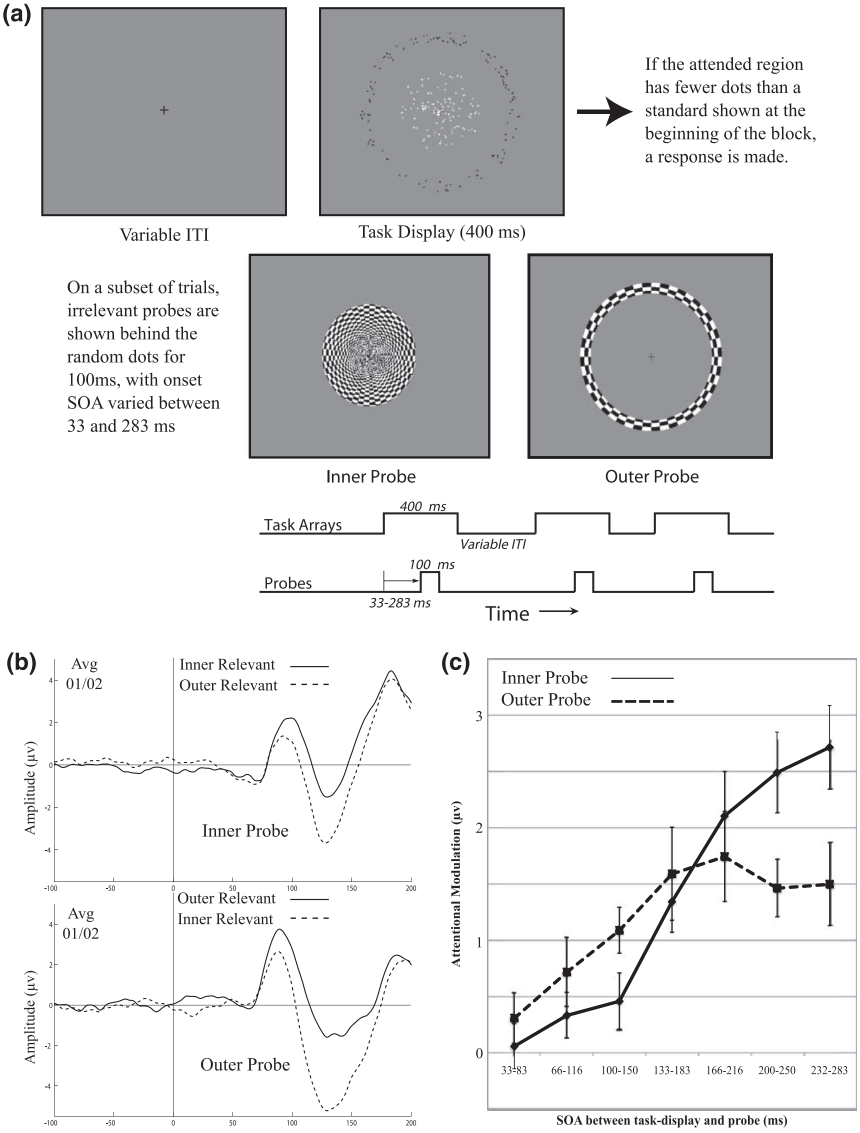
Carly J. Leonard and Steven J. Luck
Center for Mind and Brain, University of California, Davis, CA, USA

When gaze is directed towards a relevant object during natural vision, the "attentional window" must be adjusted around this object to filter out surrounding irrelevant visual input. This adjustment must be made dynamically, because the retinal size of a visual target is not necessarily predictable, depending on the object's size, its viewing distance, and occlusion. This concept of the adjustment of an attentional window within which sensory information is facilitated bears much resemblance to the spotlight view of attention (e.g., Eriksen & St. James, 1986). However, understanding the nature and time course of the expansion and contraction of attention around the point of fixation has received relatively little study, even though it is one of the most common uses of attention in natural vision. The present study used event-related potentials (ERPs) to assess the attentional modulation of sensory processing of stimuli inside versus outside of the window and the time course over which the window is adjusted.

Preceding the adjustment of the attentional window, an initial parallel stage of feature-based detection can provide guidance about the spatial location of a potentially relevant target (Folk, Leber, & Egeth, 2002; Leonard & Egeth, 2008; Serences & Boynton, 2007). For example, the memory that a lost friend was wearing orange could help guide spatial attention to a location at which this feature is detected, allowing for further examination of a potentially relevant person in the crowd. To improve the ability to discriminate friend from stranger, the attentional window must be adjusted around the currently examined object. Our task is much like this, requiring participants to initially attend to a spatially broad region, detect a region containing relevant information, and then adjust the spatial window of attention accordingly to facilitate target discrimination.

We used the occipital P1 ERP component (onset at 70–80 ms) to measure sensory processing. Previous studies show that P1 amplitude is increased for both target and nontarget stimuli when they appear in attended locations compared to when they appear in unattended locations, indicating a modulation of sensory processing (Heinze, Luck, Mangun, &

Please address all correspondence to Carly J. Leonard, Center for Mind & Brain, 267 Cousteau Place, Davis, CA 95616, USA. E-mail: cjleonard@ucdavis.edu



If the attended region has fewer dots than a standard shown at the beginning of the block, a response is made.

On a subset of trials, irrelevant probes are shown behind the random dots for 100ms, with onset SOA varied between 33 and 283 ms

Figure 1. (a) Example of trial, probes, and timeline; in actual displays, random dots were red and blue on middle grey background. (b) ERPs to the inner and outer probes, when they occur in the task-relevant and task-irrelevant region. (c) The time course of attention effects to the inner and outer probes.

Hillyard, 1990). For example, after attention has shifted to a target item in a visual search array, a small square can be flashed at either the target location or a nontarget location, and this flash will receive enhanced sensory processing (evoking a larger P1 response) if it appears at the target location rather than at a nontarget location (Luck, Fan, & Hillyard, 1993).

Whereas previous P1 studies have examined extrafoveal attention, the present study manipulated whether attention was directed to a narrow region at fixation or a broader annular region around fixation. The SOA between task display and probe was varied to examine the temporal evolution of attention. The sensory response to a probe will be modulated if, when it onsets, spatial attention is differentially allocated between task-relevant and task-irrelevant locations. At early SOAs, when the spatial window of attention has not yet been adjusted, the probe-evoked P1 should not differ depending on whether it appeared at the task-relevant or irrelevant region on that trial. The onset of spatial selection should be measurable as the earliest SOA at which modulation of the probe-elicited P1 is observed.

METHODS

ERPs were recorded from 12 participants while they performed a numerosity judgement task on an array of random dots (Figure 1a). Participants attended a single target colour (red or blue) for each 5-minute block, with the relevant colour equally likely to occur at the inner and outer region of the display. Thus, the observers did not know whether the inner or outer region would be task relevant until stimulus onset, and the spatial window of attention was presumably adjusted once the observer perceived which region contained the attended colour. On two-thirds of trials, a 100 ms probe (composed of black-and-white checks) was presented at either the inner or outer region of the task display, with SOA between the task array and the probe varying between 33 and 283 ms. The remaining trials contained no probe. The probe-absent waveform was subtracted from the probe-present waveform to isolate the probe-elicited ERP. Trials with EEG artifacts or eye movements were excluded.

RESULTS

Figure 1b shows the waveforms for inner and outer probe stimuli (after subtracting the response to the task display) when the task-relevant information occurred at the inner or outer location, averaged over SOAs. For both inner and outer probes, the P1 wave was larger when attention was directed to the region that was probed. This attention-related difference was measured for the inner probes and for the outer probes at each 50 ms SOA range, and the results are shown in Figure 1c. The attentional modulation of the P1 was near zero for the shortest SOA range (33–83 ms) and increased at longer delays. One-sample t-tests for each 50 ms SOA range revealed that the attention effect for the inner probe became significantly greater than 0 at the 133–183 ms SOA bin ($p < .001$) and remained significant for all subsequent bins. The attention effect for the outer probe emerged slightly earlier, reaching significance at the 66–116 ms SOA bin ($p < .001$), and also remained significant for all subsequent bins.

DISCUSSION

Overall, the results elucidate the temporal dynamics of how spatial attention adjusts to specifically select a task-relevant region of the visual field. This process occurs quickly, with attentional modulation of subsequent visual input occurring within about 100–150 ms of the onset of a relevant feature in the field. This attentional effect appeared earlier for irrelevant probes in the periphery compared to those presented foveally, although it was clearly visible for probes at both locations. These results demonstrate that adjusting the spatial extent of attention around the point of fixation leads to a change in sensory processing, just as attending to extrafoveal locations modulates sensory processing. They further demonstrate that the spatial window of attention can be adjusted rapidly on the basis of relevant features, a process critical for the coordination of goal-directed, stimulus-appropriate behaviour.

REFERENCES

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics*, *40*, 225–240.

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception and Psychophysics*, *64*, 741–753.

Heinze, H. J., Luck, S. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. I. Evidence for early selection. *Electroencephalography and Clinical Neurophysiology*, *75*, 511–527.

Leonard, C. J., & Egeth, H. E. (2008). Attentional guidance in singleton search: An examination of top-down, bottom-up, and intertrial factors. *Visual Cognition*, *16*, 1078–1091.

Luck, S. J., Fan, S., & Hillyard, S. A. (1993). Attention-related modulation of sensory-evoked brain activity in a visual search task. *Journal of Cognitive Neuroscience*, *5*, 188–195.

Serences, J. T., & Boynton, G. M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, *55*, 301–312.

Modelling effects of object naming on long-term object recognition memory

Michael L. Mack, Jennifer J. Richler, Sean Polyn, and Thomas J. Palmeri

Vanderbilt University, Nashville, TN, USA

Lupyan (2008) demonstrated that overtly naming objects leads to impaired long-term recognition memory compared to objects rated for preference (naming effect). Critically, this effect was reflected in a reduction in hit rates for named objects with no differences in false alarm rates. Participants failed to recognize previously named objects but were not biased to falsely recognize lures matched to named objects.

Lupyan proposed a representational shift account of this naming effect whereby overtly naming an object activates top-down information of the object's category that then augments the bottom-up object representation. This top-down categorical information thus distorts the representation for the named object creating a mismatch between the memory representation of the object and the perceptual representation of the object when it is presented again later during a memory test. This mismatch leads to a lower hit rate for named objects. A central tenet of the representational shift account is that the memory distortion for named objects arises from a dynamic interaction between top-down category information and bottom-up perceptual representations. This account tacitly assumes that naming objects and rating their preference produces representations of otherwise equivalent memory strength, and that any difference in memory strength would not predict the naming effect.

More recent work argues that differences in recognition memory between named objects and preference rated objects are more likely a consequence of stronger memory following a preference rating (Richler, Gauthier, & Palmeri, 2010). Rating preference of objects leads to better memory than naming because rating preference is a more effortful task that leads to stronger representations (e.g., Craik & Lockhart, 1972).

Unfortunately, both the representational shift and depth of processing accounts are merely verbal theories. The current work investigates the plausibility of both accounts within the framework of the REM model, a

---

Please address all correspondence to Michael Mack, Department of Psychology, Vanderbilt University, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240-7817, USA. E-mail: michael.mack@vanderbilt.edu

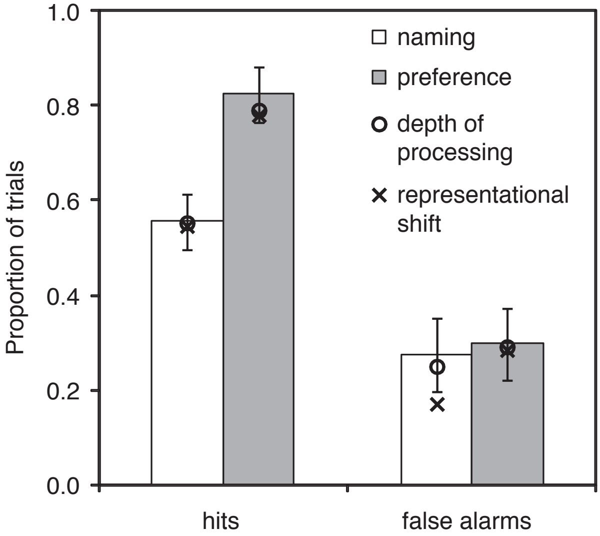


Figure 1. Model simulation results with behavioural data from Richler et al. (2010). Behavioural data shown in columns plot the proportion of hits and false alarms for objects named at study (white bars) and rated for preference (grey bars); error bars represent 95% confidence intervals. Model predictions are plotted as data points for the depth of processing model (circles) and representational shift model (crosses).

leading computational model of human recognition memory (Shiffrin & Steyvers, 1997). REM represents objects as a vector of features, with parameters that determine the value, strength, and probability of feature encoding. At test, the representation of a test object is compared to each trace in memory through calculation of a likelihood ratio. If the average of these likelihood ratios is greater than a criterion, the test object is labelled "old"; otherwise it is labelled "new".

Both the representational shift and depth of processing hypotheses can be modelled by manipulations of different mechanisms within REM. The representational shift is implemented as a postencoding shift of memory traces for named objects towards the prototypical object. The depth of processing account is modelled as a difference in the strength of encoding of feature values, with lower strength for preference versus named objects. This results in memory traces with more encoded values for rated objects than named objects.

Each of these two hypotheses was instantiated by a single parameter difference in REM between naming and preference, with all other parameters between the two encoding tasks held constant. One simulated experiment consisted of 40 study objects (20 in the naming condition, 20 in the preference condition) and 40 matched lures, just like the human experiments. Model performance was based on the average hit rates for the study objects and false alarm rates for the lures from 1000 experiment simulations. Best-fitting parameters for both models were found with the simplex method by minimizing the summed squared error between the model and behavioural data from Richler et al. (2010).

Results of the model simulations are shown in Figure 1 along with the Richler et al. (2010) behavioural results. To briefly summarize the behavioural data, the naming effect is reflected by the lower hit rate for objects named at study (white bars) relative to objects rated for preference (grey bars), with no difference between naming and preference in false alarms to matched lures. This pattern of results is accounted for by the depth of processing hypothesis (circles in Figure 1), but not by the representational shift model (crosses in Figure 1). The representational shift model predicts a lower hit rate for objects named at study; but, critically, the model also predicts fewer false alarms for lures matched to named objects.

Evaluating the two accounts of the naming effect offered by Lupyan (2008) and Richler et al. (2010) within a computational framework provides two critical results. First, the representational shift account does not predict the behavioural naming effect. Second, predictions from the depth of processing account are consistent with the behavioural naming effect. These results coupled with Richler et al. provide converging evidence that the naming effect can be explained using general principles of recognition memory, where memory differences are the result of differences in the strength of initial encoding.

REFERENCES

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: Framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

Lupyan, G. (2008). From chair to “chair”: A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General*, *137*(2), 348–369.

Richler, J. J., Gauthier, I., & Palmeri, T. J. (2010). Automaticity of basic-level categorization accounts for naming effects in visual recognition memory. *Manuscript submitted for publication.*

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM: Retrieving Effectively from Memory. *Psychonomic Bulletin and Review*, *4*(2), 145–166.

Search and destroy: Observers use an inefficient explicit feature-based inhibition strategy in visual search

Jeff Moher and Howard E. Egeth

Johns Hopkins University, Baltimore, MD, USA

When searching for a target, attentional mechanisms can help an observer find relevant items. For example, observers find a target more quickly when they have explicit foreknowledge of the location or a feature of that target. However, the value of having explicit foreknowledge of properties that do not match the target (i.e., negative information) remains unclear.

Participants can implicitly deprioritize locations in visual search (e.g., inhibition of return; Posner & Cohen, 1984). Features can also be implicitly inhibited. For example, in the distractor previewing effect (Goolsby, Grabowecky, & Suzuki, 2005) and in visual marking (Braithwaite, Humphreys, & Hodsoll, 2003), participants are slower to respond to targets presented in a colour that was previously the colour of a nontarget item.

There is some evidence that observers can explicitly deprioritize locations in visual search as well. In a recent study by Munneke, van der Stigchel, and Theeuwes (2008), when one of four potential locations was cued to indicate that it would not contain the target, participants located the actual targets more quickly than on uncued trials.

In the present studies, we examined the effects of explicit foreknowledge of negative feature information. Participants searched for a target letter ("B" or "F") among four differently coloured letters. The remaining letters were a "K" and "X" (one uppercase and one lowercase), and a lowercase "b" or "f", either compatible or incompatible with the target response (cf. Eriksen & Eriksen, 1974). On "cued" trials, participants were given a cue indicating a colour that would not match the target on the upcoming trial (e.g., "Ignore Red" validly indicates that the target will not be red). On "neutral" trials, participants received an uninformative cue ("Neutral"). These cue types were randomly intermixed.

In Experiment 1, the target letter could appear in one of four fixed locations on every trial (similar to Munneke et al., 2008). The remaining letters were randomly assigned to the remaining three locations. Participants were unable

---

Please address all correspondence to Jeff Moher, Psychological & Brain Sciences, Johns Hopkins University, 3400 N. Charles St., Ames Hall, Baltimore, MD 21218, USA. E-mail: jmoher1@jhu.edu

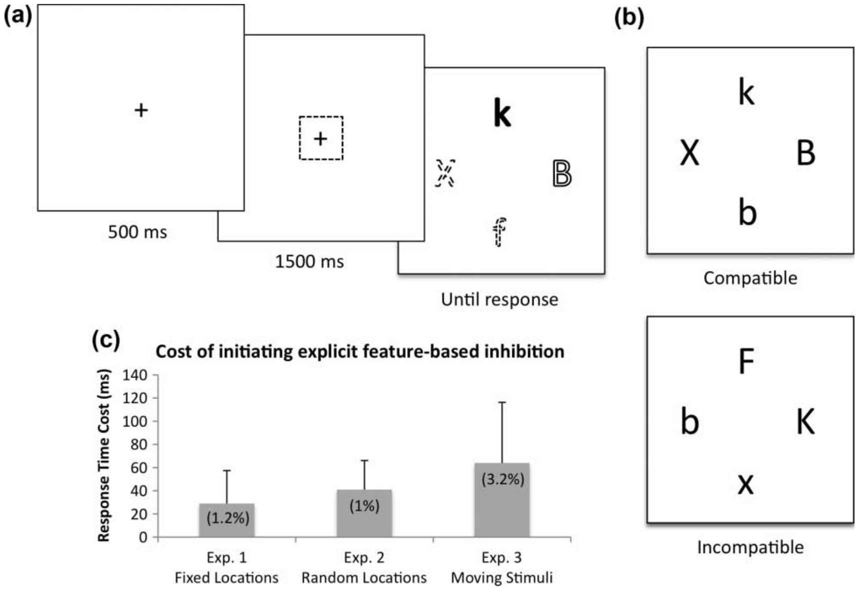


Figure 1. (A) Participants indicated the presence of an uppercase "B" or "F". In Experiment 1, letters appeared randomly in one of four fixed locations, as shown here. In Experiment 2, locations varied randomly from trial to trial. In Experiment 3, the letters moved together in a random direction on each trial. In the figure, different colours are represented by different texture patterns. On cued trials (like the one shown here), participants were cued to ignore a specific colour (represented here by an evenly dashed line). This cue validly indicated that the target would not appear in the cued colour (in this case, the distractor "f" appears in the cued colour). On neutral trials, the cue did not indicate a colour for the observer to ignore. (B) In the top image, there is a lowercase "f" that is compatible with the target response ("B"). In the lower image, there is a lowercase b which is incompatible with the target response ("F"). (C) Response time cost (cued – neutral trials) for all three experiments (accuracy cost listed in parentheses). Participants are slower to respond and make more errors in all three experiments on cued trials compared to neutral trials ($p < .05$). As location becomes less certain, the size of this cost increases.

to successfully use explicit foreknowledge of negative feature information to guide visual search; instead this foreknowledge actually slowed their search. Participants were 29 ms slower to respond on cued trials compared to neutral trials, and made more errors on cued trials (4.2%) than neutral trials (3%).

This result suggests that explicit feature-based inhibition may not be possible in visual search. However, participants could have been using an inefficient strategy of actively inhibiting the location of the to-be-ignored item after finding it, rather than applying suppression to that feature over the entire display. In Experiment 2, the four letters were randomly assigned to four of 140 possible locations, making it more difficult for the observer to anticipate where the target and distractor letters might appear. However, we still found a decrease in performance following cued trials compared to neutral trials, with an even larger decrement in performance for response time following cued trials compared to Experiment 1 (41 ms), and a similar difference in error rates (3.2% vs. 2.2%).

In Experiment 3, we attempted to eliminate any possibility that participants were using location-based mechanisms to accomplish the task. We used the same design as Experiment 1, but all items in the display were moving (together in a random direction; e.g., Andrews, Watson, Hulleman, & Braithwaite, 2010). There was an even greater cost in response time (64 ms) and errors (9.3% vs. 6.2%) for cued trials compared to neutral trials.

The results across all experiments suggest that explicit feature-based inhibition is not effective in guiding visual search. Instead, foreknowledge of negative feature information slows search (experiments conducted with visual rather than verbal cues, not reported here, produced similar results). This is consistent with an "attentional white bear" account (e.g., Tsal & Makovski, 2006) of feature-based inhibition, whereby observers attend to something they are trying to ignore. This is also consistent with results from Friedman-Hill and Wolfe (1995) showing that observers cannot limit their search to a subset of items defined by irrelevant features (i.e., "the target is among the nonred items").

Our results appear to stand in contrast to a study by Woodman and Luck (2007), who found observers searched more efficiently when the to-be-ignored colour was present (cued valid trials) than when it was not (cued invalid trials). However, we can reconcile their results with ours by suggesting that engaging in explicit feature-based inhibition results in inefficient search, but once observers initiate a top-down set to ignore a feature, they are faster to respond when the feature they are trying to ignore appears than when it does not. Therefore, rather than a "template for rejection" strategy (Woodman & Luck, 2007), we speculate that participants are engaging in a "search and destroy" strategy. They are seeking out the to-be-ignored item in order to actively inhibit it.

On cued trials in our experiments, the to-be-ignored item was a lowercase "b" or "f", either compatible or incompatible with the target. Consistent with a "search and destroy" mechanism, the compatibility effect was reduced on cued trials compared to neutral trials in all experiments, suggesting that participants were actively suppressing the identity of the to-be-ignored item. Additionally, pilot data examining all of these trial types together shows that responses are faster on cued valid trials than cued invalid trials, but fastest on neutral trials where explicit feature-based inhibition is not implemented by the observer. Further experimentation in progress is designed to more directly test for the presence of a "search and destroy" mechanism and better describe the time course of the inhibition that results from it.

Taken together, these results suggest that when participants attempt explicit feature-based inhibition, they may be using an inefficient "search and destroy" strategy that minimizes interference from the to-be-ignored item but increases overall response time and error rates. This is in contrast to explicit location-based inhibition, which observers can effectively use to decrease response times (Munneke et al., 2008). This has implications for real world search; observers performing search tasks where efficiency is important (e.g., medical screening) should not explicitly incorporate negative feature information into their search strategies.

REFERENCES

Andrews, L., Watson, D. G., Hulleman, J., & Braithwaite, J. J. (2010, May). *The cost of feature-based inhibition in dynamic search*. Poster presented at the 10th annual meeting of the Vision Sciences Society, Naples, FL.

Braithwaite, J., Humphreys, G., & Hodsoll, J. (2003). Color grouping in space and time: Evidence from negative color-based carryover effects in preview search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(4), 758–778.

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, *16*(1), 143–149.

Friedman-Hill, S., & Wolfe, J. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology*, *21*(3), 531–551.

Goolsby, B., Grabowecky, M., & Suzuki, S. (2005). Adaptive modulation of color salience contingent upon global form coding and task relevance. *Vision Research*, *45*(7), 901–930.

Munneke, J., van der Stigchel, S., & Theeuwes, J. (2008). Cueing the location of a distractor: An inhibitory mechanism of spatial attention? *Acta Psychologica*, *129*, 101–107.

Posner, M., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). Hove, UK: Lawrence Erlbaum Associates Ltd.

Tsal, Y., & Makovski, T. (2006). The attentional white bear phenomenon: The mandatory allocation of attention to expected distractor locations. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 351–63.

Woodman, G., & Luck, S. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance*, *33*(2), 363–377.

The role of incidental object fixations in repeated search: Looking *at* versus looking *for* an object in a scene

Melissa L.-H. Võ and Jeremy M. Wolfe
Harvard Medical School, Brigham and Women's Hospital, Cambridge, MA USA

Imagine baking bread in your kitchen: You reach for the flour in the shelf above you, then some yeast from the cupboard, followed by water from the tap. It seems intuitive that familiarity with a scene will speed search for different objects. For example, baking bread in your own kitchen will be much faster than baking bread in a kitchen you have never been in before.

However, previous work on repeated search (i.e., repeatedly searching for different items in an unchanging display) has shown that, despite increasing experience with a stimulus display, search efficiency does not increase as a function of search repetitions (e.g., Kunar, Flusberg, & Wolfe, 2008; Oliva, Wolfe, & Arsenio, 2004; Wolfe, Klempen, & Dahlen, 2000). Using letter displays, Kunar and colleagues (2008), for example, showed that despite searching through the same letter display for hundreds of trials, the slope of RT × Set size functions did not decrease over the course of the experiment. The authors argued that even though the stimulus display was well memorized, participants found a new visual search to be faster than accessing their memory for the display.

This is counterintuitive. Doesn't attention to distractor objects make it easier to find those objects when they become targets later in a repeated search task? Previous studies have shown that incidental fixations on objects during search can improve object recognition memory for these objects (e.g., Castelhano & Henderson, 2005; Võ, Schneider, & Matthias, 2008). Thus, at least some information from incidentally fixated objects seems to be retained in memory. In a series of three experiments, we tested whether this information can be used to benefit repeated search. That is, does looking *at* an object help us when looking *for* that object later on?

Please address all correspondence to Melissa Le-Hoa Võ, Visual Attention Lab, Harvard Medical School, 64 Sidney Street, Suite 170, Cambridge, MA 02139, USA. E-mail: mlvo@search.bwh.harvard.edu

This work was supported by the Deutsche Forschungsgemeinschaft (VO 1683/1-1) to MV and by grants to JMW (NEI EY017001, ONR N000141010278). We also want to thank Erica Kreindel for help in data acquisition.

Figure 1. (Caption on next page.)

GENERAL METHODS

In all experiments, participants searched, one after the other, for 15 different objects in an unchanging scene (see Figure 1a). A target word, presented in the centre of the screen, indicated what object to look for. Participants were instructed to search for the object as fast as possible and, once found, to press a button while fixating the object. This triggered auditory feedback. While the scene remained, a new word appeared to indicate the next search target. One block consisted of 15 searches through each of 10 different scenes (150 trials). Eye movements were recorded with an EyeLink1000 (desktop mounted; SR Research, Canada) sampling at 1000 Hz.

In Experiment 1, participants searched repeatedly for objects in the same scenes across three, 150 trial blocks. In Experiment 2, object search was preceded by a letter search task in which letters were superimposed on each object in the scene that would become a search target. Thus, participants would look *at* future target objects without looking *for* them. The subsequent two blocks were repeated search for objects as in Experiment 1. In Experiment 3, participants previewed each scene for 30s before searching for objects. They were asked if the room was most likely to be inhabited by a female or a male person.

RESULTS

There are three key findings: (1) There was no difference in search performance as a function of search epoch within the first search, all Fs $<$ 1. That is, the first 15 repeated searches through the same scene were essentially similar to each other despite increasing scene familiarity. In line with this finding, the number and duration of incidental fixations on future targets did not correlate with subsequent search speed for those items, $r = -.009$; $p > .05$. (2) There was no improvement in the time to first target fixation for the first searches (Block 1) of Experiment 2 ($M = 805$ ms, $SD = 54$) compared to Experiment 1 ($M = 823$ ms, $SD = 64$), $F < 1$, despite several seconds of scene exposure during letter search prior to object search. In Experiment 3, 30 s of scene inspection did not improve search either, but actually slowed search down compared to Experiment 1 ($M = 1014$ ms, $SD = 82$), $F(1) = 6.86$, $p < .05$. Participants basically searched the scene as if they had never seen the scene before (see Figure 1B). That is, having looked

Figure 1. A Trial sequence of Experiments 13. B Very similar heat maps of search fixation distributions here for the target object jam jars in Block 1 across Experiments 13. Warmer colours indicate longer gaze durations. C Heat maps of search fixation distributions searching for jam jars in Blocks 1, 2, and 3 of Experiment 1. Note that fixation distributions narrow down considerably from Block 1 to Block 3. [To view this figure in colour, please visit the online version of this Journal.]

at an object before did not help looking *for* it afterwards. (3) However, the second and third searches for an object were several hundred ms faster than the first search (e.g., Experiment 1: Block 1 = 823 ms vs. Block 2 = 433 vs. Block 3 = 249 ms), implying a radical benefit for having looked for that specific object previously. This can also be seen in a reduction of search space by means of fixation distributions (see Figure 1c).

DISCUSSION

This study demonstrates the different effects of looking *at* and looking *for* objects during repeated search in real-world scenes. Even though looking *at* an image is adequate to commit it (and some of its objects) to memory (e.g., Hollingworth, 2006), this experience does not alter search for those objects. Searching for other items in a scene (here, superimposed letters) had no impact on search even though the letter search forced fixation on each eventual target object. Even when given an initial task that required 30 s of engagement with the scene and its objects, the first searches for those objects did not improve. However, when an object had been previously searched for and found, the next search for that specific object was much faster—in the range of 500 ms. We conclude that, although incidental object fixations might raise recognition memory for these items above chance (see Castelhano & Henderson, 2005; Võ et al., 2008), this information is not functional for the object searches performed by our observers. Apparently, simply searching for the object *de novo* was more effective than relying on memory. However, memory of a specific search for a specific object is capable of producing very substantial speeding of subsequent search for the same object despite many intervening searches for different objects in different scenes. This raises the important question of how memory for previously searched *for* objects differs—quantitatively and qualitatively—from memory for previously looked *at* objects, and why one memory is functional for searching, and the other is not.

REFERENCES

Castelhano, M., & Henderson, J. M. (2005). Incidental visual memory for objects in scenes. *Visual Cognition*, *12*(6), 1017–1040.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 58–69.

Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2008). The role of memory and restricted context in repeated visual search. *Perception and Psychophysics*, *70*(2), 314–328.

Oliva, A., Wolfe, J. M., & Arsenio, H. C. (2004). Panoramic search: The interaction of memory and vision in search through a familiar scene. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(6), 1132–1146.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features on object search. *Psychological Review*, *113*(4), 766–786.

Võ, M. L.-H., Schneider, W. X., & Matthias, E. (2008). Transsaccadic scene memory revisited: A "theory of visual attention (TVA)" based approach to recognition memory and confidence for objects in naturalistic scenes. *Journal of Eye-Movement Research*, *2*(2), 1–13.

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Post-attentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 693–716.

Changes in ambiguous object structure are associated with shifts of attention

Yangqing Xu and Steven L. Franconeri

Northwestern University, Evanston, IL, USA

Reversible figures, such as the duck/rabbit illusion or Necker cube, produce shifts between multiple possible interpretations even while the sensory input remains unchanged. What changes in the mind and brain lead to these alternative representations?

Perceptual reversal is affected by many factors, such as expectation (Bruner & Minturn, 1955), familiarity (Rock, Hall, & Davis, 1994), and intention or will (Long & Toppino, 2004). Reversible figures may also be influenced by the location of selective attention. Cueing spatial attention to a part of the image associated more closely with one interpretation (e.g., the mouth of the duck or the rabbit) can bias observers towards that interpretation (Tsal & Kolbert, 1985). Using the Necker cube figure, another study used fMRI to show that perceptual reversals were associated with increased activity across many visual areas in the hemisphere contralateral to the "front" side of the cube. These increases were similar to those found in a control task that required observers to selectively attend to that side of the cube (Slotnick & Yantis, 2005).

The goal of the present experiment is to demonstrate this close association between perceptual switches and shifts of spatial attention at a high temporal resolution. Previous research has demonstrated that attentional shifts can be tracked by electrophysiological correlates, where potentials contralateral to an attended location are relatively more negative compared to those ipsilateral to that location (e.g., Brisson & Jolicoeur, 2007; Klaver, Talsma, Wijers, Heinze, & Mulder, 1999; Luck & Hillyard, 1994; Vogel & Machizawa, 2004). We predicted that preceding reports of a

---

Please address all correspondence to Yangqing Xu, Northwestern University, 2029 Sheridan Rd, Evanston, IL 60208, USA. E-mail: xuy@u.northwestern.edu

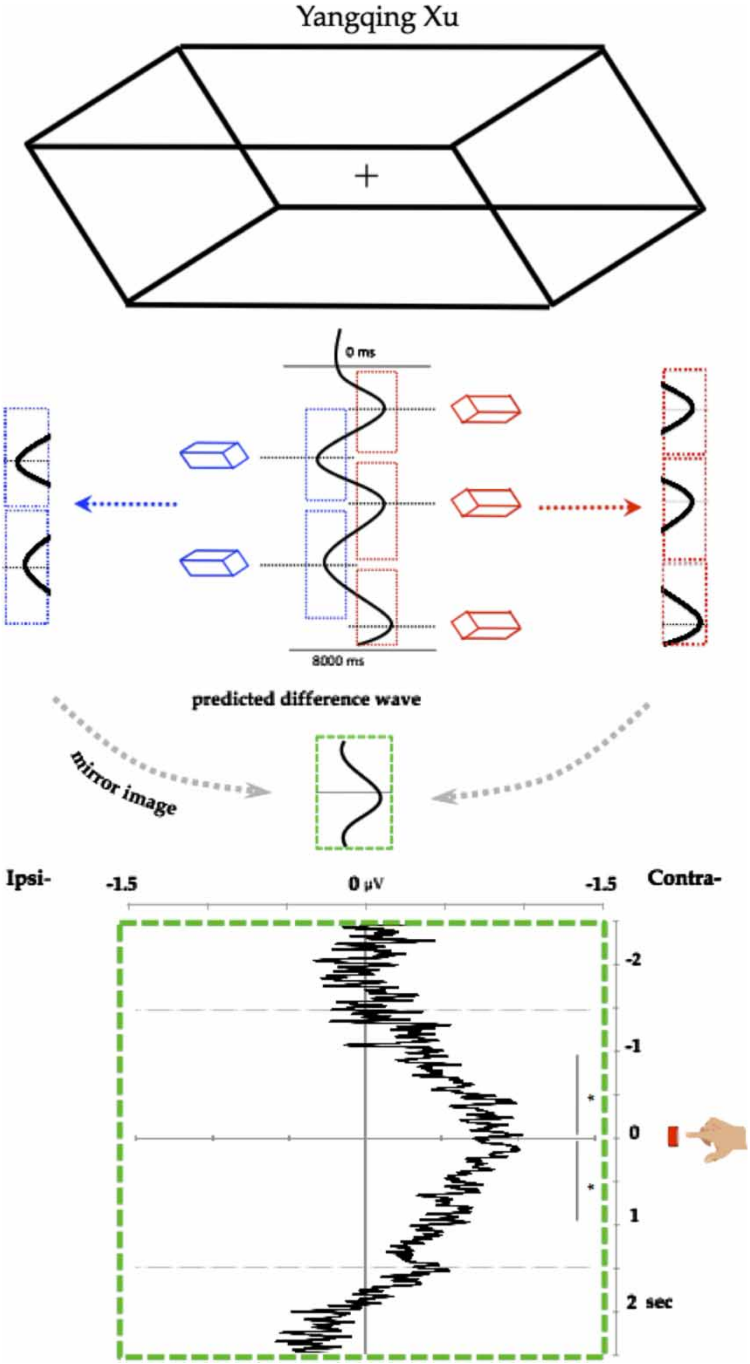


Figure 1. (Caption on next page.)

perceptual switch, there would be relative negativity on the electrode sites contralateral to the perceived "front" side of the cube.

A total of 11 participants with normal or corrected-to-normal vision completed the experiment. The ambiguous display was a modified version of a Necker cube (see Figure 1). The image consisted of two squares, which are rotated 30° counterclockwise, connected by four horizontal lines, all drawn with black lines against a white background. The length of the edges of the diamonds was 4° in visual angle and the distance between the closest corner of the diamond and the fixation was 2° in visual angle. The experiment was run using SR research Experiment Builder on Windows XP. All stimuli were displayed on a 17-inch ViewSonic E70fB CRT monitor with 1024 × 768 pixel resolution and a 60 Hz refresh rate. The view distance was approximately 55–60 cm.

Before staring the experiment, all subjects were given fixation training using a flickering pattern that "jumps" when fixation is broken (Guzman-Martinez, Leung, Franconeri, Grabowecky, & Suzuki, 2009). Participants pressed a button on the game pad to initiate a trial. Each trial began with a 1600–2000 ms fixation display, to minimize the impact of previous trials on the EEG signal. The ambiguous cube was then displayed for 8 s, during which the participants were asked to press a corresponding button each time their percept changed, while at all times maintaining fixation. The entire experiment lasted approximately 120 minutes, including ERP cap preparation, breaks, and task practice.

ERP was recorded using a Biosemi Active II EEG/ERP system. All sites were re-referenced to the postrecording average of the left and right mastoids and low-pass filtered at 80 Hz. We recorded from the following sites according to the 64-channel modification of the international 10/20 system: F3/4, C3/4, PO3/4, P5/6, P7/8, PO7/8, O1/2, POz, Oz, Horizontal and Vertical EOG. Eye movements were also monitored by a table-mounted SR-Research Eyelink 1000 Remote eyetracker. Trials with eye movements were rejected by the eyetracker and recycled. Additional trials were rejected at the analysis stage due to HEOG/VEOG deviations.

The EEG data was epoched within a response-locked time window spanning 2 s before and 2 s after the report of a perceptual change, and baseline corrected to the 200 ms prestimulus period. Figure 1 depicts the

Figure 1. The top panel depicts a sample test display. The middle panel shows a schematic version of the analysis technique. Within an 8-second trial, there could be several reports of a perceptual switch in the structure of the cube. We took response-locked ERPs at each report of a switch (see methods for details), and collapsed the two types of percept reports into a difference wave showing activity contralateral to the new perceived front of the cube. The bottom panel presents the grand average of this difference wave across subjects, showing more PO7/PO8 negativity contralateral to the front face of the cube before and after the switch report, suggesting a shift of attention toward the new front side. [To view this figure in colour, please visit the online version of this Journal.]

difference waveforms as the subtraction of the ipsilateral waveforms (PO7 for the left-in-front percept and PO8 for the right-in-front percept) from the contralateral waveforms (PO7 for the right-in-front percept and PO8 for the left-in-front percept). The data were analysed in a 2 × 16 repeated measures ANOVA in which the factors were electrode (contralateral or ipsilateral to the perceived "front" side"), and 16 measurement time windows of 250 ms (from 2 s before to 2 s after the response). There was a main effect of time, $F(15,\ 150) = 5.1$, $p < .001$, reflecting global potential changes unrelated to the contralateral differences of interest. There was a significant main effect of electrode, $F(1,\ 10) = 9.3$, $p = .012$, reflecting more negativity contralateral to the perceived "front" side. There was an Electrode × Time interaction, $F(15,\ 150) = 2.0$, $p = .021$, reflecting more negativity at electrode sites contralateral to the front of the cube at time bins from 1000 ms before to 1000 ms after switch responses, all $t\text{s} > 2.7$, $p\text{s} < .023$, but not time bins before or after that range. However, substantial individual differences among subjects led to violations of sphericity assumptions, and the interaction did not survive a Greenhouse-Geisser correction, $p = .137$. These results suggest that participants attended to the front side of the figure at least 1000 ms before they reported that side as being in front.

In summary, the present study demonstrated a close correlation between shifts of spatial attention and perceptual reversals. Selective attention may increase the cortical response associated with that area, which may bias the perception of the corresponding configuration (Slotnick & Yantis, 2005), and might help maintain one interpretation once it is formed. These results suggest that instead of merely enhancing the processing of a currently selected location, selective attention may play a role in altering the perceived spatial structure of an object. The present findings also have implications for other forms of bistable perception, such as ambiguous depth from motion and apparent motion, where visual structure may also be related to the distribution of attention across space and time.

REFERENCES

Brisson, B., & Jolicoeur, P. (2007). A psychological refractory period in access to visual short-term memory and the deployment of visual-spatial attention: Multitasking processing deficits revealed by event-related potentials. *Psychophysiology*, *44*, 323–333.

Bruner, J. S., & Minturn, A. L. (1955). Perceptual identification and perceptual organization. *Journal of General Psychology*, *53*, 21–28.

Guzman-Martinez, E., Leung, P., Franconeri, S. L., Grabowecky, M., & Suzuki, S. (2009). Rapid eye-fixation training without eye tracking. *Psychonomic Bulletin and Review*, *16*, 491–496.

Klaver, P., Talsma, D., Wijers, A. A., Heinze, H. J., & Mulder, G. (1999). An event-related brain potential correlate of visual short-term memory. *NeuroReport*, *10*, 2001–2005.

Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin*, *130*(5), 748–768.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*, 291–308.

Rock, I., Hall, S., & Davis, J. (1994). Why do ambiguous figures reverse? *Acta Psychologica*, *87*(1), 33–57.

Slotnick, S. D., & Yantis, S. (2005). Common neural substrates for the control and effects of visual attention and perceptual bistability. *Cognitive Brain Research*, *24*(1), 97–108.

Tsal, Y., & Kolbert, L. (1985). Disambiguating ambiguous figures by selective attention. *Quarterly Journal of Experimental Psychology*, *37A*, 25–37.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*, 748–751.