Visual Cognition

Object Perception, Attention, and Memory 2007 Conference Report 15th Annual Meeting, Long Beach, California, USA

Organized by; Monica Castelhano; Steve Franconeri; Kim Curby; Sarah Shomstein

PLEASE SCROLL DOWN FOR ARTICLE

Ψ Psychology Press
Taylor & Francis Group

Object Perception, Attention, and Memory 2007 Conference Report 15th Annual Meeting, Long Beach, California, USA

Organized by Monica Castelhano, Steve Franconeri, Kim Curby, and Sarah Shomstein

Context affects figure–ground organization via perceptual grouping

Joseph L. Brooks and Jon Driver

Institute of Cognitive Neuroscience, University College London, London, UK

Gestalt psychology inspired wide interest in perceptual grouping and figure–ground organization (FGO). But while often mentioned together in textbooks, these topics have rarely been directly related. Whereas FGO determines the assignment of edges to regions, grouping involves linking otherwise discontinuous image elements. Here, we address whether FGO might be affected by perceptual grouping. Specifically, we asked whether FGO of an ambiguous edge can be affected by a distant, but perceptually grouped, unambiguous edge-segment elsewhere in the image. To address this question we created displays with three critical sections (Figure 1A). The biased section contained an edge with clear FGO based on a local figure–ground cue. We used edge-region grouping by common fate motion (Palmer & Brooks, in press) in which the texture in one region bounding an edge moved in common fate with the edge, whereas texture in the other region did not, causing edge assignment to the grouped region. The neutral section of the display contained an edge with no local figure–ground bias. Whether the biased and neutral sections were connected to one another was ambiguous because of a rectangular occluder between them.

To test whether FGO in the neutral section could be affected by the biased context and to determine the extent to which this interaction depended on grouping with the context, we varied the grouping relationship between the two sections of the display. We did this in two ways: Edge-grouping and

Address correspondence to Joseph L. Brooks, Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK. E-mail: joseph.brooks@ucl.ac.uk
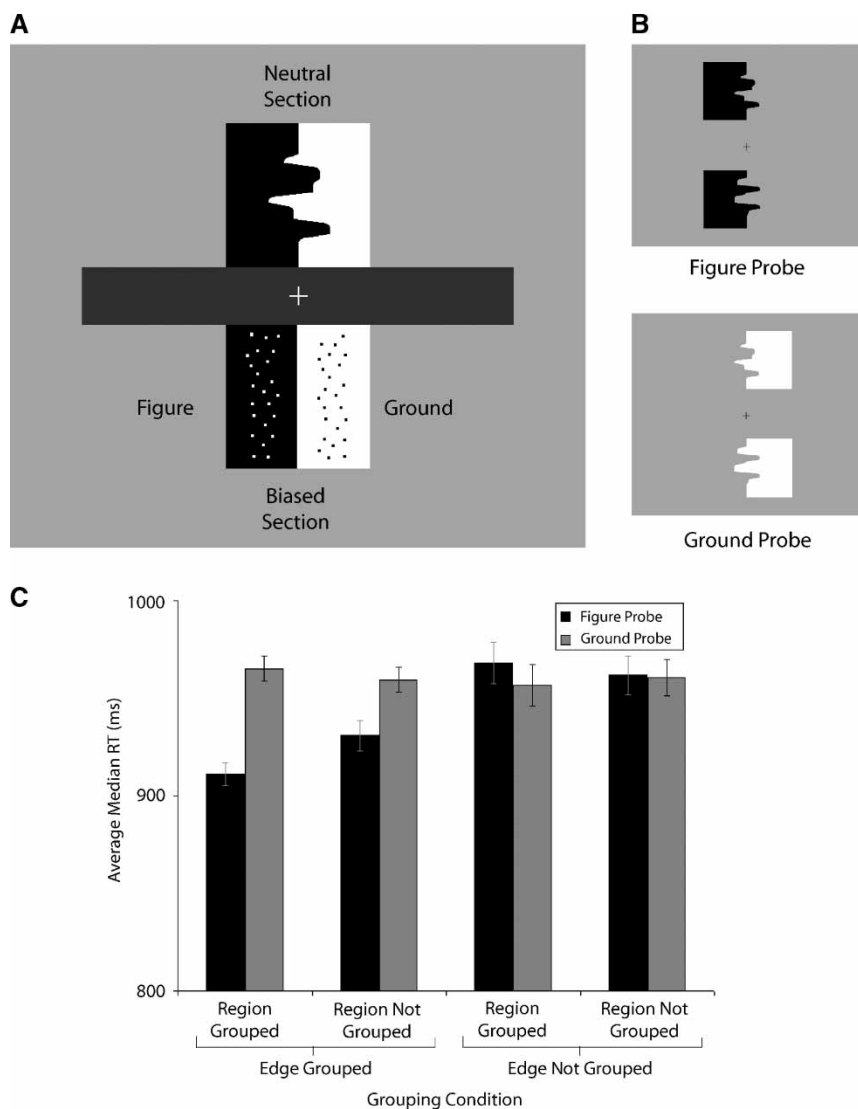
Figure 1. (A) An annotated example of the displays in Experiment 1 and prime displays in Experiment 2. The text labels for neutral and biased regions and the figural assignment of the biased region were not present in the experimental displays. (B) Examples of figure and ground probe displays (assuming that figure–ground organization in the biased region is as shown) used in Experiment 2. The shapes above fixation in each probe display are edge-matches to the prime display shown in Figure 1A. (C) Average median reaction time to figure and ground probes as a function of edge and region grouping conditions in Experiment 2. Error bars represent 95% within-subjects confidence intervals (Loftus & Masson, 1994).

region-grouping. The edges in the two sections always oscillated back-and-forth. Edge-grouping was present when edges in the two sections were collinear and moved together at the same speed, suggesting that they may be part of the same edge. Edge not-grouped conditions were characterized by edges that moved separately at different speeds in the neutral and biased sections, suggesting that they may not be perceptually related to one another. We manipulated region-grouping by varying the colour similarity of the regions in the two sections. When the colours were the same, there was region-grouping between the biased and neutral sections. There was no region-grouping when the colours were different.

EXPERIMENT 1

In Experiment 1 we assessed FGO in the neutral section with a subjective report task. That is, we asked participants which region (left or right of the edge) appeared to be figural. We found effects of edge-grouping, $F(1, 11) = 59.76$, $p < .0001$; region-grouping, $F(1, 11) = 8.16$, $p < .016$; and their interaction, $F(1, 11) = 9.40$, $p < .011$, on reports of figure–ground organization in the neutral region. In conditions with both edge-grouping and region-grouping between the biased and neutral sections, participants reported FGO in the neutral region to be consistent with that in the biased region 80% of the time, significantly greater than 50% unbiased reporting, $t(1, 11) = 12.36$, $p < .001$. When there was edge-grouping but no region-grouping, the context effect was significantly reduced to 71%, $F(1, 11) = 42.07$, $p < .0001$, but still significantly greater than 50%, $t(1, 11) = 8.36$, $p < .001$. Region-grouping on its own, however, was not sufficient to raise context-consistent judgements above 50%, $t(1, 11) = 1.13$, $p < .284$.

EXPERIMENT 2

In Experiment 2 we used an objective task to measure figure–ground organization. In the short-term edge-matching (STEM) task (Driver & Baylis, 1996), participants saw a 1.5 s prime display like those used in Experiment 1. The prime display was followed by one of two types of probe displays (visible until response; Figure 1B), each containing two shapes with curvy edges. One of the edges was the same as the edge in the neutral section of the prime display, and participants were instructed to select that edge. However, whereas in figure probe test displays the figure–ground assignment of this edge was the same as in the prime displays, it was reversed in ground probe displays. The difference in perceived shape along the edge between the prime and probe edges in ground probe displays causes slower and less accurate edge-matching responses than for figure probe displays for which

the perceived shape along the edge is the same in prime and probe displays. This difference between figure and ground probe edge-matching performance provides an indirect measure of FGO.

Participants were significantly slower (Figure 1C) and less accurate in selecting ground probes than figure probes in the condition with both edge and region grouping between the neutral and biased sections; accuracy (fig probe = 5.5% vs. ground probe = 10.0% errors), $F(1, 15) = 63.77$, $p < .0001$; reaction time, $F(1, 15) = 122.91$, $p < .0001$, as well as in the condition with edge grouping alone; accuracy (figure probe = 8.5% vs. ground probe = 10.3% errors), $F(1, 15) = 22.61$, $p < .0001$; reaction time, $F(1, 15) = 32.01$, $p < .0001$. Figure probe and ground probe reaction times and accuracy were not significantly different in conditions without edge grouping, i.e., region-grouped only and no-grouping conditions. This suggests that when the edges in the neutral and biased sections were grouped, figure–ground organization in the neutral section was affected by that of the biased region. This influence was then manifested in performance on the STEM task. As in Experiment 1, the context effect due to edge-grouping was weaker when the regions were not grouped. Furthermore, region-grouping alone did not cause a significant effect of context in the STEM task.

DISCUSSION

These results suggest that FGO and perceptual grouping have more than a textbook association under the rubric of Gestalt psychology. Rather, perceptual grouping plays an important role in maintaining consistent FGO along edges, especially in cases where continuity of image edges is disrupted by factors such as occlusion. Edge context should be added to the list of cues to FGO used by the human visual system and it may prove useful in computer vision applications. Most importantly, models of perceptual organization in human vision must be updated to account for effects of grouping on figure–ground organization. Palmer, Brooks, and Nelson (2003) have suggested that grouping is ubiquitous in visual processing. Our results provide new evidence for this hypothesis.

REFERENCES

Driver, J., & Baylis, G. C. (1996). Edge-assignment and figure–ground segmentation in short-term visual matching. *Cognitive Psychology*, *31*(3), 248–306.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review*, *1*(4), 476–490.

Palmer, S. E., & Brooks, J. L. (in press). Edge-region grouping in figure–ground organization and depth perception. *Journal of Experimental Psychology: Human Perception and Performance*.

Palmer, S. E., Brooks, J. L., & Nelson, R. (2003). When does grouping happen? *Acta Psychologica*, *114*(3), 311–330.

Factors governing inhibition of occluded regions in superimposed objects

Destinee Chambers and Kyle Cave

Department of Psychology, University of Massachusetts, Amherst, MA, USA

While previous studies debated whether visual attention is guided by location-based or object-based mechanisms, more recent research investigates the interaction between these two. Overlapping objects have been used from the beginning of research on object-based attention. They are effective stimuli for evoking object/location interactions in laboratory tasks because they share a space but require the perception of multiple objects. Understanding how overlapping objects are separated should provide insight about how location-based selection is shaped by object boundaries.

Some studies show that when a target object is partially occluded by a distractor object, there is less attention at the overlapping region than at visible regions of the target (Cave & Wake, 2001; Davis & Driver, 1997; Moore & Fulton, 2005). In other circumstances, though, the occluded region receives as much attention as the visible target regions (Haimson & Behrmann, 2001). Understanding the factors that manipulate this occluder inhibition (OI) will help to understand the object/location interaction. We have conducted a series of experiments to identify those factors.

We recently examined the role of figural complexity and whether real-world objects (RWOs) differed from abstract geometric objects in their ability to elicit OI. Cave and Wake (2001) found OI using RWOs with detailed patterns within the object contours (Figure 1A), while Haimson and Behrmann (2001) had much simpler stimuli and found no OI. We found OI in both the figurally complex condition with the patterned interiors (Figure 1C) and the noncomplex condition with the unpatterned interiors. These results, accompanied by our replication of Haimson and Behrmann (Figure 1B), eliminated figural complexity as a controlling factor and confirmed that OI was not specific to RWOs.

Address correspondence to Destinee Chambers, Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA. E-mail: destineechambers@gmail.com
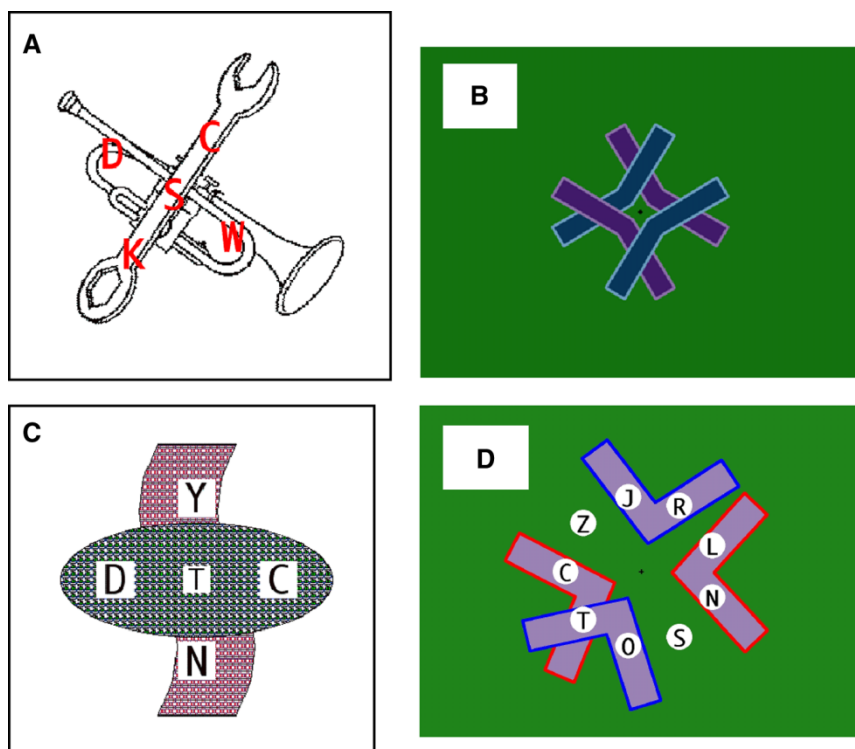
Figure 1. Sample stimuli. (A) Cave and Wake (2001); (B) Haimson and Behrmann (2001) replication—overlapping objects were alternating colours of purple or blue; (C) figural complexity experiment (patterned condition)—each pattern was made up of at least two different colours; (D) object number experiment. To view this figure in colour, please see the online issue of the Journal.

Specificity to verbal cueing paradigms was a possible factor because both Cave and Wake (2001) and the present experiment described found OI without using spatial cues. Instead, subjects received verbal instructions for which object to attend. We deduce that OI is not specific to verbal cueing paradigms because both Moore and Fulton (2005) and Davis and Driver (1997) also observed OI using spatial cues.

In both Cave and Wake (2001) and the figural complexity experiment, the overlapping region was in the centre of the display and OI appeared, while Haimson and Behrmann (2001) had more peripheral regions of overlap and no OI. This suggested that position of the occlusion might be an important factor. However, this idea can be eliminated by comparing Moore and Fulton (2005) with Cave and Wake (2001) and the figural complexity experiment.

We also know that binocular disparity and motion (suggested by Haimson & Behrmann, 2001, and Moore & Fulton, 2005, to explain why

their results differed from Davis & Driver, 1997) are not necessary for OI because all of the studies in our laboratory showed OI using static monocular displays.

Our recent experiments tested whether the number of objects or the number of overlapping regions affected OI. The results show that these are also unlikely factors.

NUMBER OF OBJECTS

Haimson and Behrmann (2001) used four objects and did not find OI; that's more objects than any study showing OI (Cave & Wake, 2001; Davis & Driver, 1997; Moore & Fulton, 2005). This experiment tested whether OI only appears with a small number of objects. Perhaps the computations necessary to identify occlusions become a burden when many objects are present.

Stimulus

There was a ring of boomerang-shaped objects centred around fixation. In any given trial, the participant saw two to five objects. Only two objects overlapped. The border of each was red or blue. (See Figure 1D.)

Task

After the objects appeared, one end of one of the overlapping objects was briefly cued to indicate that it was the target. Then nine letter probes appeared briefly in a circle around fixation, with one at the region of overlap. Participants reported any four of the letters and the colour of the target.

Results

Accuracy reporting the letter at the occluded region was lower when the target was occluded than when it was unoccluded, demonstrating OI. Results did not vary with number of objects, showing that OI is independent of display size.

LEVEL OF TARGET OBJECT PROCESSING

All studies that were able to show OI (excluding Moore & Fulton, 2005, and Davis & Driver, 1997) required that subjects report some characteristic of

the target (identity or colour). Removing this extra task might limit the processing of the target, which could eliminate OI.

Stimulus and task

These were the same as the object number experiment, except that participants did not report the colour of the target.

Results

Removing the need to report a property of the target eliminated OI. OI depends on the degree to which the target is processed.

Why did Moore and Fulton (2005) find OI, even though their subjects were not required to report anything about the target? Their display includes object motion, which may make the object more salient and cause it to be processed more fully than it would be otherwise. Moving one object while others are stationary may have an effect similar to requiring that the object colour be reported.

DISCUSSION

The first experiment shows that OI can occur with any number of objects, demonstrating that object number is unimportant in eliciting OI. Comparison across the two experiments shows that OI is not specific to pure discrimination tasks, because the last experiment used a pure discrimination task and produced no OI. Also, comparing the two suggests that the number of overlaps is not a determining factor, because these two experiments both only have one area of overlap and yield different results. However, we cannot conclusively say that OI can be observed with four or more overlaps present.

Comparing our recent results with previous results suggests that the degree to which the target object is processed determines whether occluded regions are inhibited. Our next experiment will attempt to induce OI by using motion to increase target salience. The results illustrate that the degree to which attention modifies visual processing depends on perceptual goals.

REFERENCES

Cave, K. R., & Wake, D. (2001). *Attending to locations of superimposed visual objects*. Paper presented at the annual meeting of the Psychonomic Society, Orlando, FL, November.

Davis, G., & Driver, J. (1997). Spreading of visual attention to modally versus amodally completed regions. *Psychological Science*, *8*, 275–281.

Haimson, C., & Behrmann, M. (2001). Cued visual attention does not distinguish between occluded and occluding objects. *Psychonomic Bulletin and Review*, 8(3), 496–503.
Moore, C. M., & Fulton, C. (2005). The spread of attention to hidden portions of occluded surfaces. *Psychonomic Bulletin and Review*, 12(2), 301–306.

Gender aftereffects in face silhouettes reveal face-specific mechanisms

Nicolas Davidenko

Department of Psychology, Stanford University, Stanford, CA, USA

Nathan Witthoft and Jonathan Winawer

Department of Psychology, Stanford University, Stanford, CA, and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Recently a parameterized face space has been created using profile face silhouettes (Davidenko, 2007). Face silhouettes provide enough information for accurate judgements of age, gender, attractiveness, and race, and their parameterization allows us to characterize the physical factors that affect these judgements. Here we use these stimuli to further probe the representation of face gender, by implementing a novel rapid, implicit adaptation paradigm. Using this paradigm, we first show that gender aftereffects like those found with textured front-view faces (e.g., Webster, Kaping, Mizokami, & Duhamel, 2004) occur with face silhouettes. Furthermore, these aftereffects transfer across changes in contrast polarity and left–right orientation of the adapting silhouettes, neither of which affect the perception of a silhouette as a face; however, vertically inverting the adapting stimuli greatly reduces aftereffects. Finally, we show that adapting to silhouettes alters gender judgements of front-view faces, and vice versa, suggesting that gender processing in front-view faces and silhouettes relies on common neural mechanisms, and that these are the site of the adaptation effect. We conclude that face silhouettes provide a useful tool to test hypotheses about face space representation.

Address correspondence to Nicolas Davidenko, Department of Psychology, 450 Serra Mall, Building 420, Stanford, CA 94107, USA. E-mail: ndaviden@psych.stanford.edu

METHOD: RAPID IMPLICIT ADAPTATION

The stimuli were eight male and eight female parameterized "adapting" face silhouettes, and one gender-neutral "target" face silhouette (see Davidenko, 2007). Participants completed a one-page questionnaire consisting of nine face silhouettes (Figure 1A). The first eight adapting silhouettes were either all female or all male, and the ninth silhouette was always the same gender-neutral target. The first eight (adapting) silhouettes were rated on attractiveness, race, or age, and only the ninth (target) silhouette was rated on gender. We refer to this as "rapid, implicit adaptation" because participants completed the questionnaire in about 1 minute and were not instructed to
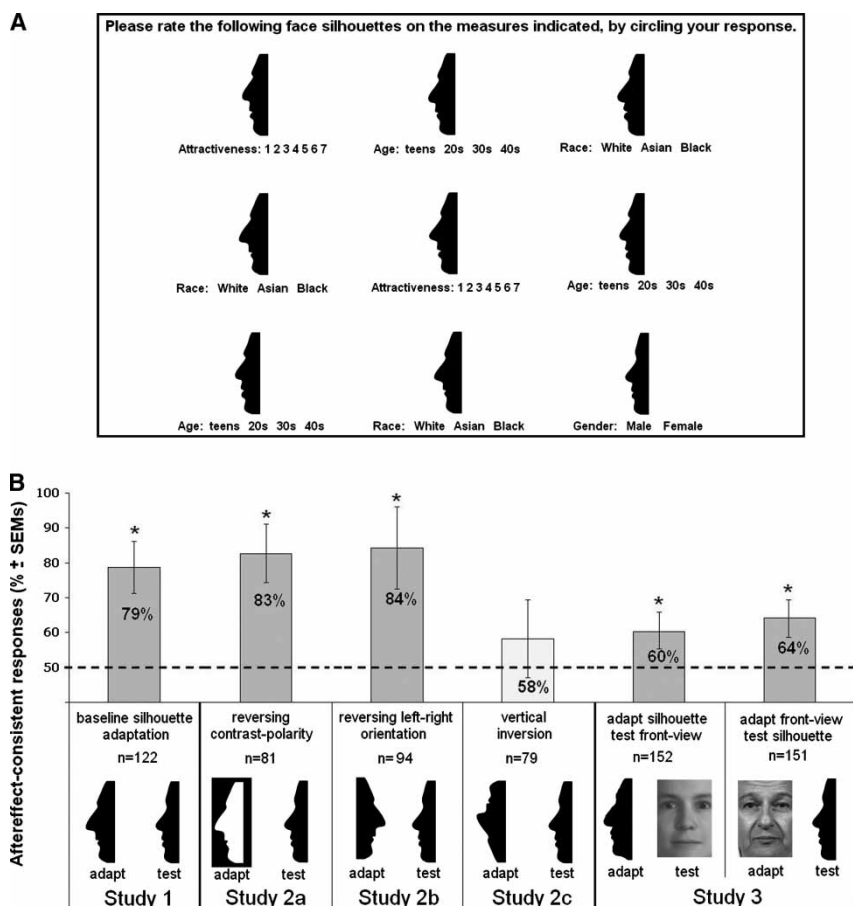
Figure 1.   (A) A sample questionnaire used in Study 1. (B) The summary of results from Studies 1, 2, and 3 as the percentage of aftereffect-consistent responses compared to chance level.

attend to gender of the eight adapting stimuli; they simply provided nine ratings on the silhouettes.

STUDY 1: GENDER AFTEREFFECTS

One hundred and twenty-two participants were assigned to either the adapt-female or the adapt-male condition. The response to the target silhouette was the variable of interest. Only 2 of the 59 adapt-female participants rated the target silhouette as "female", compared to 39 of the 63 adapt-male participants, $\chi^2(1) = 47$, $p < .001$, equivalent to 79% aftereffect-consistent ratings (Figure 1B). Gender aftereffects thus occur with face silhouettes, and they can be elicited in this rapid, implicit adaptation paradigm.

Do the observed aftereffects reflect adaptation to faces per se, or can they be explained by adaptation to low-level image properties?

STUDY 2A: REVERSING CONTRAST POLARITY

Reversing the contrast polarity of a silhouette does not obviously alter the interpretation of the stimulus as a face despite reversing the contrast of all local contours. This study used the same procedure and stimuli as in Study 1, except that the eight adapting face silhouettes were white-on-black. The gender-neutral target remained black-on-white. Only 4 of the 42 adapt-female participants, compared to 29 of the 39 adapt-male participants, rated the target silhouette as "female", $\chi^2(1) = 35$, $p < .001$, equivalent to 83% aftereffect-consistent ratings. Thus, gender aftereffects in silhouettes transfer across changes in contrast polarity.

STUDY 2B: REVERSING LEFT–RIGHT ORIENTATION

Next we considered the possibility that the gender aftereffects could be explained by local shape or curvature adaptation (see Suzuki & Cavanagh, 1998). To reduce the contribution of shape adaptation, we flipped the eight adapting face silhouettes so that they faced right (while the target remained facing left). Only 8 of the 50 adapt-female participants, compared to 37 of the 44 adapt-male participants, rated the target as "female", $\chi^2(1) = 18$, $p < .001$, equivalent to 84% aftereffect-consistent ratings, suggesting that the gender aftereffects cannot be explained by low-level shape adaptation alone.

STUDY 2C: VERTICAL INVERSION

As with face photographs, vertical inversion of face silhouettes impairs face processing (Davidenko, 2007; Yin, 1969). We reasoned that if gender aftereffects depend on face-specific processing, inverting the adapting stimuli should reduce aftereffects. Indeed, with inverted adapting silhouettes, 19 of the 36 adapt-female participants, compared to 29 of the 42 adapt-male participants, rated the upright target as "female", showing no significant effect of adapting condition, $\chi^2(1) = 1.8$, $p > .15$, equivalent to 58% aftereffect-consistent ratings.

STUDY 3: TRANSFER ACROSS IMAGE FORMAT

The results so far suggest that gender adaptation is invariant to transformations that preserve the face percept. To test this hypothesis further, we measured whether aftereffects transfer between profile face silhouettes and greyscale front-view faces. Since most image properties are different between these two image formats, any preservation of aftereffects would demonstrate that the aftereffects operate on a high, face-specific level of visual representation. The procedure for this study was the same as in the previous studies, but with four conditions: Either the adapting stimuli were silhouettes (152 participants, either *adapt-female or adapt-male*) with a gender-neutral, greyscale, front-view target face constructed with the Face Modeler software, or the adapting stimuli were the greyscale front-view counterparts of these silhouettes and the target stimulus was the gender-neutral face silhouette used in the previous studies (151 participants). Remarkably, gender aftereffects persisted across these drastic changes in face image format. In the adapt-silhouette conditions, the proportion of aftereffect-consistent ratings of the front-view target was 60%, $\chi^2(1) = 5.7$, $p < .02$, and in the adapt-front-view conditions, the proportion of aftereffect-consistent ratings was 64%, $\chi^2(1) = 11.6$, $p < .001$.

Our results suggest that gender processing and adaptation occur, at least in part, at a high level of visual representation that is relatively insensitive to image transformations that preserve the face percept. Corroborating recent fMRI evidence that face silhouettes elicit selective activity in the fusiform face area (Davidenko, Remus, Glover, & Grill-Spector, 2007), we have shown evidence that gender processing in face silhouettes shares common mechanisms with gender processing in front-view faces. This provides further validation of the face silhouette methodology, and suggests that parameterized face silhouettes can contribute to our understanding of the dynamics of gender representation and adaptation in the general face domain.

REFERENCES

Davidenko, N. (2007). Silhouetted face profiles: A new methodology for face perception research. *Journal of Vision*, *7*(4): 6, 1–17.

Davidenko, N., Remus, D., Glover, G., & Grill-Spector, K. (2007). *Sensitivity to face format and distinctiveness in face-selective cortex*. Abstract to be presented at the 2007 annual meeting of the Society for Neuroscience.

Suzuki, S., & Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1315–1341.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*, 557–561.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.

Object-based storage in visual working memory and the visual hierarchy

Tao Gao

Zhejiang University, Hangzhou, Zhejiang Province, P. R. China, and Yale University, New Haven, CT, USA

Mowei Shen, Zaifeng Gao, and Jie Li

Zhejiang University, Hangzhou, Zhejiang Province, P. R. China

How visual information is represented in visual working memory has been extensively studied in the past decade. Luck and Vogel (1997) proposed a "strong object" hypothesis, suggesting that the capacity of working memory is limited by the number of objects, regardless of the complexity of each object. However, other studies showed that the capacity drops and the object-based benefits in storage disappear when the memory items become complex (Alvarez & Cavanagh, 2004; Olson & Jiang, 2002; Xu, 2002). Olson and Jiang (2002) put forward a "weak object" hypothesis, suggesting that working memory is limited by both the number of objects and the composition of those objects. However, it is still unclear why object-based benefits are weak in working memory. Most importantly, how can theories

Address correspondence to Tao Gao, Department of Psychology, Yale University, New Haven, CT, 06520-8205, USA. E-mail: Tao.Gao@yale.edu

about visual perception, working memory, and attention be put into a coherent framework in terms of this "weak object" hypothesis?

Considering that visual working memory and visual perception are actively intertwined types of processing, we suggested that theories on how visual information is extracted at multiple levels of perceptual processing can shed light on how this information is stored in working memory. Here we were particularly interested in the reversed hierarchy model of visual perception (Hochstein & Ahissar, 2002), suggesting that vision begins with the automatic and parallel processing following the bottom-up pathway of the visual hierarchy. According to this model, feedforward processing is not directly available to conscious perception. Explicit visual perception only begins when processing reaches high cortical levels and proceeds in top-down fashion to gradually encompass detailed information available at low cortical areas as needed.

To account for the "weak object" property, we hypothesized that object-based storage of working memory originates from highly discriminable information that has already been segmented as integrated objects even before entering working memory. For low discriminable information, which can only be accumulatively encompassed into the object representation via serial reentrant processing, storage is not object based. We tested this hypothesis in the current study. Unlike previous work focusing on revealing the capacity difference for storing different types of information, the current study explored whether there are dissociated mechanisms for storing visual information represented at different levels of the reversed visual hierarchy. In particular, we asked whether information represented at a certain level of the visual hierarchy can be stored as integrated objects in working memory.

Experiments 1 and 2 explored the object-based working memory by investigating the selection of visual information with different degrees of discriminability. Provided visual objects are selected into working memory as integrated whole, then when only one feature of an object is required to be stored in working memory, the other task irrelevant features from the same object will also be automatically selected. This object-based automatic selection was probed by observing whether changes in the task-irrelevant feature dimension can distract change detection on the target feature dimension. In Experiment 1, memory materials were coloured shapes (Figure 1). Observes were instructed to detect change from one basic feature (e.g., colour) while ignoring changes from another task-irrelevant feature (e.g., shape). The results showed that an irrelevant feature change can significantly impaired change detection in the target feature dimension,
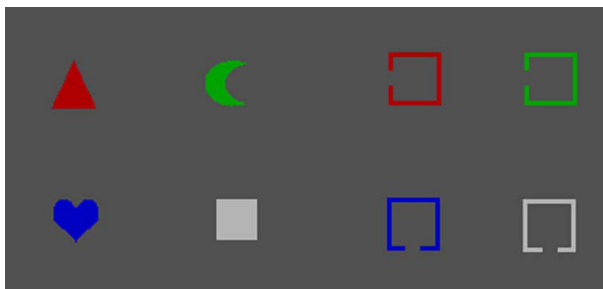
Figure 1. The memory materials displayed in Experiment 1 (left) and Experiment 2 (right). In the real experiments, each item in the memory array has a distinct colour (e.g., white, black, red, green). In Experiment 1, observers retain the colours or the shapes of the memory array. In Experiment 2, observers retain the colours or the gap's orientations. To view this figure in colour, please see the online issue of the Journal.

indicating an automatic object-based storage of highly discriminable basic features.

Experiment 2 explored whether similar object-based automatic storage can extend to low discriminable information. Memory items were coloured rings with a small gap on the frame (Figure 1). Visual search studies have shown that gap orientation is difficult to discriminate, requiring serial attentive processing (Woodman & Luck, 2003). Our results revealed that, as in Experiment 1, irrelevant colour change can distract detecting changes in the orientation gap. However, irrelevant change in gap orientation has no detrimental effect on detecting colour change. Thus, there is evidence suggesting that information requiring attentive processing is not automatically selected into working memory.

Experiment 3 and 4 further explored object-based working memory by estimating the efficiency of consolidating information of high and low discriminability. For information consolidated in working memory as integrated objects, different dimensions of information along with their conjunctions should be consolidated with identical efficiency. Results of Experiment 3 confirm this hypothesis, showing identical efficiency for consolidating colour, shape, and colour-shape conjunction. However, results of Experiment 4 showed that the consolidation of gap orientation is much slower than that for colour. These results suggest that whereas different dimensions of highly discriminable information are consolidated as integrated wholes, there is a dissociated process for consolidating low discriminable information.

Experiments 5 and 6 explored object-based storage by investigating the maintenance of bound visual information with different discriminabilities. Provided individual features are stored as integrate objects in visual working memory, conjunction information should be successfully stored as long as individual features are retained. Therefore, there should be no binding-specific decrement in change detection performance (Wheeler & Treisman, 2002). Results of Experiment 5 showed that the colour-shape binding could be maintained as well as individual features. However, results of Experiment 6 showed that performance detecting changes from the binding of colour-orientation was significantly worse than that of detecting individual colour change or gap orientation change. These results indicate that whereas different dimensions of highly discriminable information are maintained as integrated objects, highly and low discriminable information are dissociable when stored in working memory.

Based on these results, we proposed that storage in visual working memory is not a unitary process. Rather, there are dissociable processes for storing visual information extracted at different stages of visual perception. The mapping between processes in working memory and visual perception indicates that how visual information is extracted during visual perception directly influences how this information will be selected, consolidated, and maintained in visual working memory. Object-based storage in working memory is distinct from the fully perceived object representation assembled by reentrant attentive processing and operates over highly discriminable information, which has already been integrated as the output of feed-forward, parallel processing.

REFERENCES

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*, 106–111.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*, 791–804.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong object" hypothesis. *Perception and Psychophysics*, *64*, 1055–1067.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48–64.

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 121–138.

Xu, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 458–468.

Segregating targets and nontargets in depth eliminates inhibition of nontargets in multiple object tracking

Harry H. Haladjian, Carlos Montemayor, and
Zenon W. Pylyshyn

Center for Cognitive Science, Rutgers University, Piscataway, NJ, USA

Multiple Object Tracking (MOT) is a useful paradigm for studying properties of visual attention. In a typical MOT task, eight or more identical objects are presented on a computer screen. Several of these objects are distinguished as targets by flashing briefly at the beginning of a trial. These objects then move in a random and unpredictable manner, and the observer is asked to identify the targets at the conclusion of this movement (as described in Pylyshyn, 2001; Pylyshyn & Storm, 1988). Observers can easily track four or five objects among identical distractors in a varying range of conditions. Such results have been interpreted as the function of visual indexes (also called FINSTs), which are the preattentive individuation mechanisms proposed by Visual Indexing Theory (Pylyshyn, 2001).

In a recent MOT study, we showed that target-target pairs tend to be confused more often than target-nontarget pairs. We argued that this may be due to the inhibition of nontargets. This interpretation was supported by a study that used probe dots (as developed by Watson & Humphreys, 1997) and showed that nontargets were indeed inhibited (Pylyshyn, 2006). Inhibition of irrelevant objects is a general phenomenon and several varieties of attentional inhibition in vision have been reported (Theeuwes & Godljn, 2002; Tipper, 2001), including the Inhibition of Return (IOR). Such inhibition appears to be object-centred, with IOR having a significant object-centred component (Tipper, Driver, & Wheeler, 1991). Yet, there is much that we do not know about visual inhibition. For example, there has been some discussion in the literature about whether inhibition is applied strategically in a top-down fashion when required by the task at hand (Watson & Humphreys, 1997) or whether objects are automatically marked for inhibition in a bottom-up manner (Theeuwes, Kramer, & Atchley, 2001).

The inhibition of nontargets raises special problems for Visual Indexing Theory. In our studies, the nontargets themselves appear to be inhibited

Address correspondence to Harry H. Haladjian, Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854-8020, USA. E-mail: haladjian@ruccs.rutgers.edu

rather than everything in the display that is not a target, such as the empty space between the objects. Previous experiments demonstrated the object-based nature of nontarget inhibition by comparing probe detection performance when the probe appeared on targets, nontargets, and locations in the empty space between and near objects, and found that only the nontargets were inhibited (Pylyshyn, 2006). If inhibition can occur on multiple moving nontargets and not on the space through which they move, the question arises: By what mechanism can the inhibition be associated with the moving nontargets? The only mechanisms we have proposed for tracking objects are visual indexes, which hypothetically are already being fully used to track targets. This challenge was also noted in the context of visual search among moving items where nontargets showed inhibition (Ogawa, Takeda, & Yagi, 2002).

In a recent study, we used probe detection during MOT to show that moving nontargets are inhibited while identical but stationary nontargets are not inhibited (Pylyshyn, 2006). We also measured inhibition on moving nontargets that were identical in appearance to the targets (nontarget circles) compared with those that were different from the targets (nontarget squares). We found that moving square nontargets and moving circle nontargets were inhibited equally (Pylyshyn, 2006). These data showed that static nontargets (which are easily distinguished from moving targets) are not inhibited as predicted by the top-down hypothesis, but moving nontargets that clearly differ in shape from targets (e.g., squares instead of circles) are still inhibited. This result indicates either that all unattended moving objects were inhibited or that the difference between square and circular nontargets was not sufficient to enable the segmentation of the different-shaped nontargets from potentially interfering with the tracking task. Perhaps if the featural difference was one that could be detected preattentively, nontargets could be segregated from targets and it would be unnecessary to inhibit them in order to keep them distinct from targets in the MOT task.

The present experiment further explores the inhibition process by examining the conditions under which observers are able to keep targets distinct from nontargets during a 3-D MOT task on a stereoscopic computer display. This display was designed to create the appearance of objects moving on two separate depth planes. Since there is evidence that depth can be used effectively to separate objects in a preattentive manner (Nakayama & Silverman, 1986), our experiment addresses whether identical moving nontargets on a different depth plane are sufficiently different or separable from targets so that inhibition of these nontargets is unnecessary. The pursuit of this question may elucidate the role of inhibition in keeping targets distinct from nontargets during MOT and whether inhibition applies only when nontargets are not preattentively separable from targets.
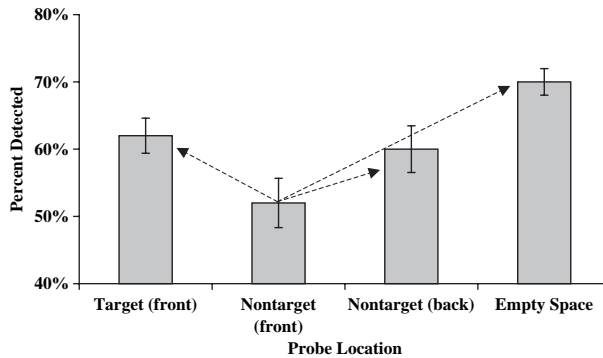
Figure 1.   Probe detection performance during MOT on two depth planes. Note: Arrows indicate significant differences in probe detection performance ($p < .05$).

Subjects in this experiment were tested for the ability to perceive the two different depth planes through stereoscopic glasses (only these data were analysed, $n = 15$). The 80 trials consisted of twelve identical circles, eight appearing on the front plane and four appearing on the back plane. Four objects were identified as targets at the start of the trial by flashing briefly. These targets always appeared on the front plane so that there would be four targets and four nontargets on the front plane and four nontargets on the back plane. The subjects were asked to track the targets and look for a probe dot that might appear anywhere on the screen, including on target and nontarget items. We found that probe detection on front nontargets (52% correct) was significantly lower than probe detection on front targets (62%), back nontargets (60%), and empty space (70%). No other significant differences were found between these conditions. (See Figure 1.)

The results from this experiment indicate that nontargets on the same plane as targets are inhibited, while nontargets on a different depth plane are not inhibited. This suggests that nontargets appearing on a different depth plane from targets are preattentively separable and thus do not require inhibition, while nontargets on the same plane as targets do require inhibition. This study supports the hypothesis that inhibition functions in a bottom-up manner and is manifest only when nontargets could be confused with targets.

REFERENCES

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*(6059), 264–265.

Ogawa, H., Takeda, Y., & Yagi, A. (2002). Inhibitory tagging on randomly moving objects. *Psychological Science*, *13*(2), 125–129.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*(1–2), 127–158.

Pylyshyn, Z. W. (2006). Some puzzling findings in multiple object tracking (MOT): II. Inhibition of moving nontargets. *Visual Cognition*, *14*(2), 175–198.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

Theeuwes, J., & Godijn, R. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception and Psychophysics*, *64*(5), 764–770.

Theeuwes, J., Kramer, A. F., & Atchley, P. (2001). Spatial attention in early vision. *Acta Psychologica*, *108*(1), 1–20.

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *Quarterly Journal of Experimental Psychology*, *54A*(2), 321–343.

Tipper, S. P, Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology*, *43A*(2), 289–298.

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review*, *104*(1), 90–122.

Target and distractor processing in visual search: Decomposition of the N2pc

Clayton Hickey, Vincent Di Lollo, and John J. McDonald

Simon Fraser University, Vancouver, Canada

In laboratory studies visual search has been used extensively to investigate the factors that govern attentional selectivity. In a typical visual-search experiment observers are presented with displays containing a number of items and are asked to respond based on characteristics of a target. Human electrophysiological studies of search have identified a discrete component of the event-related potential (ERP) that appears to reflect the allocation of attention. Referred to as the N2pc, this component is a negative ERP difference beginning ~ 175 ms poststimulus at electrode sites contralateral to the target relative to electrode sites ipsilateral to the target.

In a seminal study on the topic, Luck and Hillyard (1994) suggested that the N2pc reflected neural activity involved in the suppression of distractor stimuli. Several key pieces of evidence were provided for this hypothesis. Targets that elicited the N2pc in the presence of distractors were shown to fail to elicit the N2pc when distractors were absent (Exp. 3), when distractors

Address correspondence to: Clayton Hickey, Department of Cognitive Psychology, Vrije Universiteit Amsterdam, van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands. E-mail: c.hickey@psy.vu.nl

provided essential information about the presence or absence of the target and thus could not be suppressed (Exp. 2), and when distractors shared task-relevant features with the target (Exp. 4). The distractor-suppression hypothesis has been supported and extended by a number of subsequent ERP studies.

Empirical studies from fields other than human electrophysiology, however, suggest that selection involves multiple cognitive mechanisms. For example, single-unit studies in animals have suggested that attention acts on target representation, enhancing neural activity associated with a stimulus when attention is deployed to its location. The types of enhancement identified in this work have ranged from raw increases in baseline rates of neural firing to more complicated mechanisms such as the modulation of single-neuron response selectivity to target characteristics (see Treue, 2001, for a review). Other work has challenged the idea that distractor suppression is itself a unitary phenomenon. Some studies suggest the presence of a target-centred, spatially-circumscribed suppressive field in visual search (e.g., Hopf et al., 2006), while others have shown that target selection can result in suppressive effects at locations far from the target (e.g., Cepeda, Cave, Bichot, & Kim, 1998).

The idea that that selection occurs through the action of more than one cognitive mechanism raises the possibility that lateralized ERP activity associated with attentional selection—the N2pc—may reflect more than one such mechanism. The present study was designed to investigate this hypothesis. We conducted four ERP experiments designed to isolate lateralized brain activity associated with target and distractor processing. In these experiments participants were presented with visual search displays and required to discriminate a characteristic of the target stimulus. The visual search displays had three key characteristics. First, they contained only two stimuli, one target and one distractor. The stimuli were a bright green patch, which could have either diamond or square form, and a red line, which could be either long or short. Second, the brightness of the red line was psychophysically matched by participants with that of the display background. This minimized lateralized sensory ERP activity elicited by the green line in critical experimental conditions. Third, each stimulus was presented on the vertical meridian on some trials such that lateralized ERP activity could be attributed to the other, lateralized stimulus. This isolation technique relies on the fact that stimuli presented on the vertical meridian affect brain activity in each cortical hemisphere equally, and thus do not create lateralized potentials such as the N2pc (Woodman & Luck, 2003; see also Hickey, McDonald, & Theeuwes, 2006).

For the sake of brevity, we discuss here the results from one prototypical experiment in this series. In half of this experiment participants ($N = 12$) were instructed to attend to the red line and ignore the green patch (see

Figure 1a), while in the remaining half they attended to the green patch and ignored the red line (see Figure 1d). Displays across these two conditions were thus identical in nature, allowing for differences in the ERP data to be unequivocally linked to cognitive phenomena. Critical experimental results were elicited when participants were presented with displays containing a bright green patch on the vertical meridian and a red line at a lateral screen position. When participants attended to the red line in these displays, a contralateral negativity—the N2pc—became apparent (Figure 1b). In contrast, when participants attended to the green patch in these displays (see Figure 1d), a contralateral positivity became apparent (Figure 1e). This latter positivity appears to stem from processing of the distractor stimulus. As such, we have named it the distractor positivity or P_D.
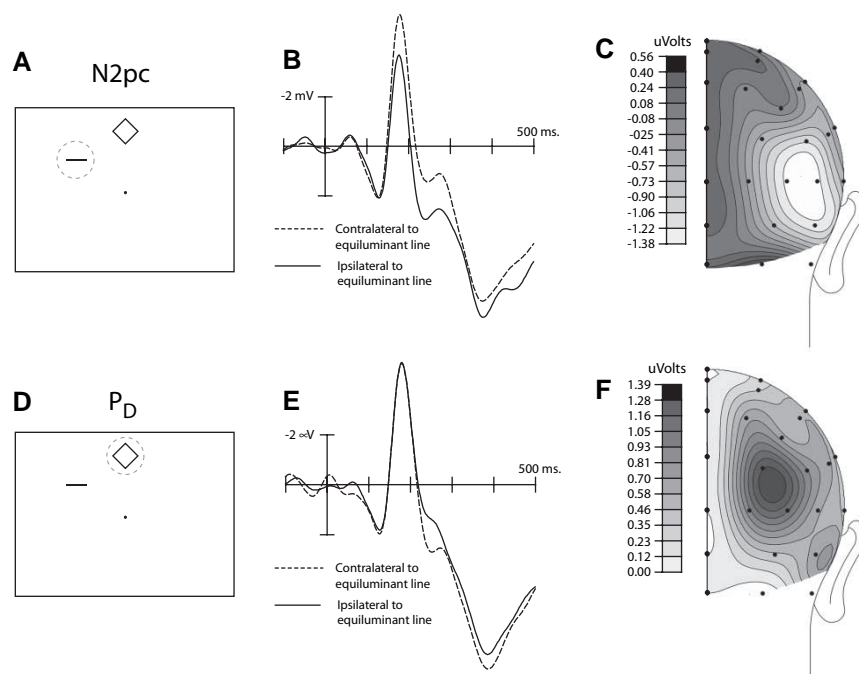
Figure 1. Stimuli exemplars and results. Top row corresponds to "attend line" condition; bottom row to "attend patch" condition. Stimuli arrays are presented in left column with broken circle denoting the focus of attention. ERPs elicited in the two conditions are presented in the middle column. Time 0 indexes stimulus onset and negative is plotted upwards. The N2pc is apparent in the upper waveform ∼175–300 ms poststimulus, while the P_D is apparent in the lower waveform ∼220–280 ms poststimulus. Half-head topographical maps of the two components are presented in the right column. These are based on spherical-spline interpolation of contralateral-minus-ipsilateral ERP activity. Black circles denote electrode positions.

We believe that the P_D indexes a mechanism of distractor suppression that acts on the cortical representation of distractor stimuli. The positive polarity of the component is consistent with this idea; inhibitory postsynaptic potentials (PSPs) in pyramidal neurons responsible for representing the distractor may result in positive charge at the scalp. In contrast, our results suggest that the N2pc is tied to processing of the target. This raises the possibility that the N2pc reflects direct processing of the target, perhaps enhancement of the cortical representation of attended stimuli. The negative polarity of the N2pc is consistent with this hypothesis; excitatory PSPs in pyramidal neurons responsible for representing the target may result in a negative charge at the scalp. Importantly, however, results tying the N2pc to the location of the target are not necessarily inconsistent with a distractor-suppression role for the component. The N2pc may reflect a mechanism that acts to shelter target representation not by stopping the output of information from cortex responsible for representing the distractor, but rather by stopping the input of distractor information into cortex responsible for target representation. As this process could occur at cortical locations responsible for target representation, it could result in an N2pc elicited contralaterally to the target.

In earlier electrophysiological studies of search attentional processing tied to target stimuli was not distinguished from attentional processing tied to distractor stimuli. The current study constitutes the first electrophysiological dissociation of target and distractor processing, and provides a novel tool for psychological investigation of attentional suppression.

REFERENCES

Cepeda, N. J., Cave, K. R., Bichot, N. P., & Kim, M. S. (1998). Spatial selection via feature-driven inhibition of distractor locations. *Perception and Psychophysics*, *60*, 727–746.

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience*, *18*(4), 604–613.

Hopf, J. M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H. J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Science*, *103*(4), 3532–3540.

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000–1014.

Treue, S. (2001). The neural correlates of attention in primate visual cortex. *Trends in Neuroscience*, *24*(5), 295–300.

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 121–138.

Evidence for separate development of working memory capacity for objects and for features in infants

Melissa M. Kibbe and Alan M. Leslie

Center for Cognitive Science, Rutgers University, Piscataway, NJ, USA

Previous research has shown that visual working memory (WM) capacity increases over the first year of life (Leslie & Kaldy, 2007; Ross-Sheehy, Oakes, & Luck, 2003). Yet many studies that have looked at infants' ability to keep track of objects have found that, though infants are limited in the number of objects of which they can keep track through space and time (Feigenson & Carey, 2003; Feigenson, Carey, & Hauser, 2002), this limit of three items does not appear to change between 6 and 12 months of age. This apparent contradiction in findings can be resolved if one considers the possibility of two distinct WM limits, one for objects themselves and one for features of those objects. Recent research with adults has indicated that, in addition to a four-item limit, the information load of the items themselves affects capacity; the greater the information load, the fewer objects can be encoded in WM (Alvarez & Cavanagh, 2004). Further, different areas of the brain have been shown to be involved in WM for a fixed number of objects and WM for object features, the former constrained by number and the latter by object complexity (Xu & Chun, 2006). The current study begins to explore whether distinct WM capacities develop differentially in infants.

While previous work has assessed infants' developing WM capacity for *individuated* objects, it has confounded this with WM for object *identities,* which are defined by the features of objects. For example, Kaldy and Leslie (2003) tested infants' WM for object identities by hiding two objects sequentially behind two screens; this method allowed them to probe memory for individual objects separately. They concluded that 9-month-old infants could remember the identities of up to two objects by testing their memory for the object that was hidden first (that is, the harder to remember of the two items). However, because they tested infants' memory for the first hidden object by showing them the last hidden object (now unexpectedly behind the first screen), in order to succeed infants need only remember that the first hidden object was *different* from the last hidden object, but not the *identity,* the specific shape, of the first hidden object. This "subsampling"

---

confound makes it difficult to pinpoint the underlying constraints on infants' developing WM capacity.

Because recent work with adults has indicated that object complexity constrains WM capacity, we sought to tease apart WM for objects and WM for features using a method after Kaldy and Leslie (2003, 2005). We used both the two-screen violation-of-expectation method and an extended version to eliminate subsampling confounds and to better define the capacity of WM in two age groups: 6-month-olds and 9-month-olds. The extended technique introduces a third shape, which allows us to test infants' WM for the identities of two objects while eliminating subsampling confounds.

EXPERIMENT 1

We sought to explicitly examine the apparent contradiction between object tracking estimates and estimates of WM by asking whether 6-month-olds could remember the existence of two objects, even though they have been shown to only remember the *identity* of one object (Kaldy & Leslie, 2005). We hid two shapes (a triangle and a disc) sequentially behind two different screens. The screen that occluded the first-hidden object was then taken away to reveal no object. Infants looked significantly longer at this outcome, suggesting that they were able to remember that two objects were hidden, though they could not necessarily remember anything about the object in question.

The results of Experiment 1 indicate that, though 6-month-olds have been shown to have WM capacity of only one for object *identity*, they are able to remember the existence of at least two objects.

EXPERIMENT 2

We asked whether 9-month-olds could remember the identities of two objects when subsampling confounds were eliminated. We hid three shapes (a triangle, a disc, and a square) sequentially behind three different screens. The screen that occluded the second-hidden object was taken away to reveal either the last-hidden object (after Kaldy & Leslie, 2003) or the first-hidden object. We found that 9-month-olds looked longer when the last-hidden object was revealed but did not look longer when the first-hidden object was revealed. This suggests that 9-month-olds are able to remember that the second-hidden object is *different* from the last object they saw (e.g., *not* triangle), but are not able to remember any identifying shape information about that object.

The results of Experiment 2 indicate that, while 9-month-olds are not able to remember the specific shape of more than one object, they are able to

remember some shape information about more than one object, namely that the second-hidden object has a shape that is different from the last object they saw. Further, it is possible that this "has a different shape" feature can apply to any object that is not the last-hidden object, regardless of the different features of those objects.

CONCLUSIONS

Taken together, these experiments provide some preliminary evidence that both 6-month-olds and 9-month-olds are able to keep track of at least two objects, but differ in their capacity to remember features of those objects. This raises the possibility that it is the ability to remember object features that develops over the first year of life and constrains WM capacity, not the ability to remember the existence of the objects themselves.

FUTURE DIRECTIONS

If indeed it is WM capacity for features that increases over the first year, what is the nature of this capacity? What is its limit? Is there a limit on number of different features, or on number of different levels of single features? Will increasing object complexity limit WM capacity for objects? We plan to explore these questions by using the method described above to test WM capacity for objects with different features (e.g., colour) and for objects with feature conjunctions (e.g., colour and shape) in three different age groups, 6-, 9-, and 12-month-olds, to ascertain the developmental trajectory of WM capacity for objects and for features.

REFERENCES

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, *6*(5), 568–584.

Fiegenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files vs. analog magnitudes. *Psychological Science*, *13*(2), 150–156.

Kaldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: Integrating "what" and "where" information. *Developmental Science*, *6*(3), 360–373.

Kaldy, Z., & Leslie, A. M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition*, *97*, 153–177.

Leslie, A. M., & Kaldy, Z. (2007). Things to remember: Limits, codes, and the development of object working memory in the first year of life. In L. M. Oakes & P. J. Bauer (Eds.), *Short- and long-term memory in infancy and early childhood: Taking the first steps toward remembering* (pp. 103–125). Oxford, UK: Oxford University Press.

Ross-sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*(6), 1807–1822.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(2), 91–95.

A cross-cultural study of shape representation

Mark D. Lescroart, Xiaomin Yue, and Irving Biederman

University of Southern California, Los Angeles, CA, USA

Jules Davidoff

Goldsmiths College, University of London, London, UK

Most current models of visual perception hold that the visual system decomposes an image of an object into a number of dimensions (e.g., colour, size, shape, etc.). In many perceptual studies using simple stimuli, "shape" is taken to be a single dimension. However, the representation of shape in our brain may be organized into a dimensional structure as well.

A number of parts-based theories of 3-D object representation (Biederman, 1987; Marr & Nishihara, 1978) hold that simple parts can be modelled as generalized cones (GCs). GCs are the volumes created by sweeping a cross-section along an axis, as when a circle is moved along a straight axis to produce a cylinder. Different volumes can be produced through variations along independent GC dimensions, such as axis curvature, the size variation of the cross-section during its sweep, and aspect ratio.

It is one thing to show mathematically, as did Marr and Nishihara (1978), that any shape can be created by GCs. But do the dimensions that define GCs have psychological reality? Given a set of stimuli generated by variations in GC dimensions, does the perception of these stimuli reflect these dimensions?

There is strong evidence that in both humans and laboratory macaques GC dimensions are indeed coded independently. Stankiewicz (2002) reported that the discrimination of noisy variations in one GC dimension by human subjects could be performed independently of the noise level on another GC dimension. Kayaert, Biederman, Op de Beeck, and Vogels (2005) showed that 95% of the variance of the firing of macaque IT cells to 2-D shapes

Address correspondence to Mark D. Lescroart, University of Southern California, Los Angeles, CA 90089, USA. E-mail: mark.lescroart@usc.edu

could be accounted for by independent representation of the GC dimensions. However, both the human subjects and the laboratory monkeys were raised in environments full of geometrically simple, regular objects, where GC dimensions could be readily appreciated. Would people raised in an environment with minimal exposure to simple, regularly shaped objects show the same effects? That is, is the sensitivity to GC dimensions observed thus far a function of robust statistics that would characterize any natural environment, or is it merely a consequence of familiarity with simple, regular objects?

The Himba, a seminomadic people in a remote area of northwestern Namibia, have little exposure to developed-world artefacts. We used a texture-segregation task to determine whether the Himba would show independent representation of two generalized-cone metric dimensions: The degree of axis curvature and the aspect ratio of a set of geons.

The task is illustrated in Figure 1. Subjects viewed a 5×5 "texture" array of curved cylinders. Each display had four different texture elements (resembling macaronis), defined by two values on each of two metric dimensions: Axis curvature (slightly- and highly-curved) and aspect ratio (wide and narrow). The four elements were (1) narrow, highly curved cylinders, (2) wide, highly curved cylinders, (3) narrow, slightly curved cylinders, and (4) wide, slightly curved cylinders. Each display was divided into two regions, each with two types of cylinders. Subjects judged, as quickly and as accurately as possible, whether the boundary between the two regions, which could lie between the second or third row or column, was vertical or horizontal.

There were three possible ways to define the boundary: (1) By axis curvature (highly curved vs. slightly curved), (2) by aspect ratio (wide vs. narrow), or (3) by a combination of the aspect ratio and axis curvature (narrow/highly curved and wide/slightly curved on one side vs. narrow/
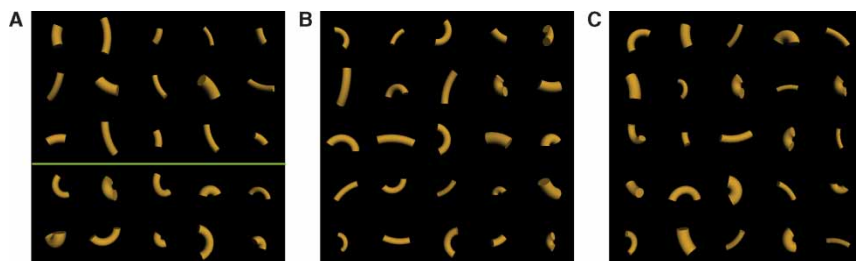

Figure 1. "Texture" arrays of curved cylinders. The divides are based on (A) axis curvature (horizontal divide—between rows 3 and 4—slightly curved on top, highly curved on the bottom [divide shown]), (B) aspect ratio (vertical divide—between columns 3 and 4), and (C) conjunction of the two dimensions (vertical again—between columns 2 and 3).

slightly curved and wide/highly curved on the other). In each of the first two conditions, subjects could perform the task based on only one shape dimension; in the third condition, they had to use information from two dimensions simultaneously (a conjunction task). To reduce the reliance on low-level cues of orientation and size, the orientation was varied randomly over 360° in 2-D and 45° in 3-D and absolute size was varied randomly by 33%.

For both Westerners and the Himba, when the boundary was defined by a single dimension (either aspect ratio or axis curvature), performance was markedly superior to when the boundary was defined by a conjunction of dimensions. Error rates for the Westerners were 2.1%, 5.9%, and 17.8%, for the axis curvature, aspect ratio, and conjunction conditions, respectively. For the Himba these values were 15.0%, 15.0%, and 32.4%, respectively. The conditions differed reliably, $F(2, 20) = 41.76$, $P_{rep} > .99$, $\eta_p^2 = .81$, as did the difference between the two groups, $P_{rep} > .99$, but importantly there was no reliable interaction between group and condition, $F(2, 20) = 1.24$, $P_{rep} < .63$, $\eta_p^2 = .11$. RTs followed a similar pattern, with longer RTs to the conjunction displays than to the single dimension displays for both groups.

Because all three displays contain exactly the same four elements with two of the elements on either side of the border, there is nothing in the displays themselves that would necessitate the difficulty of determining the boundary in the conjunction condition. It is only by our imposition of GC dimensions that the difficulty can be understood.

To verify that the low-level cues of luminance and orientation could not be the source of the difference between the conditions, we created a classifier that based its response on only orientation and intensity information. We used the feature channels of the Itti and Koch (2000) saliency map model (as published standard approximations of early visual filtering) to compute local orientation and intensity. The classifier compared the mean and variance of the distributions of intensity and orientation on each side of each possible divide. The classifier chose the divide that gave the greatest difference as its response.

The classifier performed very poorly on the task (near chance), and furthermore did not produce a human-like pattern of results (i.e., higher error rates for the conjunction condition and low error rates for the other two). Consequently, neither low-level differences in pixel intensity nor differences in orientation could explain the ordering of conditions.

The sensitivity of both the Himba and USC students to underlying dimensions of generalized cones suggests that such sensitivity does not require immersion in a regular, manufactured environment but, instead, is likely to be a consequence of robust statistics that characterize virtually any visual environment.

REFERENCES

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506.

Kayaert, G., Biederman, I., Op de Beeck, H. P., & Vogels, R. (2005). Tuning for shape dimensions in macaque inferior temporal cortex. *European Journal of Neuroscience*, *22*(1), 212–224.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society of London. *Series B: Biological Sciences*, *200*(1140), 269–294.

Stankiewicz, B. J. (2002). Empirical evidence for independent dimensions in the visual representation of three-dimensional shape. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 913–932.

Conceptual masking: Is concept the key, or does layout play a role?

Kristin O. Michod and Helene Intraub

University of Delaware, Newark, DE, USA

This research examines the role of layout in conceptual masking. Sequential pictures presented at rates that mimic or exceed normal visual scanning, are poorly remembered. Potter (1976) argued that pictures are identified within 100 ms of viewing, but require an additional 300 ms of consolidation to be remembered. The onset of a new, meaningful picture elicits processing, drawing attention from the previous picture (conceptual masking; Potter, 1976). If the preceding picture is not consolidated in memory before that time, it will be forgotten. Conceptual masking is thought to result from an inability to disengage from the onset of a novel and meaningful stimulus. Accordingly, observers can ignore "nonsense" pictures containing object-like contours, and colours but not obvious meaning, or a single meaningful picture that repeats throughout the sequence (Intraub, 1984).

Although these results are consistent with the idea that the critical feature drawing attention is conceptual change, it is important to recognize that

Address correspondence to Kristin O. Michod, Department of Psychology, University of Delaware, Newark, DE 19716, USA. E-mail: kristin@psych.udel.edu

when a new picture was presented, in addition to a new gist the picture contained a new layout. We sought to determine if the interference caused by the to-be-ignored pictures (conceptual masks) in previous studies was only elicited by the onset of a new gist or if the onset of a new layout, even when gist repeats, could also interrupt consolidation of previous pictures. We approached this question in two different ways. In Experiment 1 the conceptual masks always contained a new gist, but layout was either changed or repeated. Conversely, in Experiment 2 the conceptual masks always contained the same gist but, again, layout either changed or remained constant.

HYPOTHESES

If conceptual masking is caused solely by gist, then changes in gist (Experiment 1) will cause conceptual masking, whereas changes in layout that do not affect the gist can be ignored. Likewise, when the same gist is repeated (Experiment 2) changes in layout will not cause conceptual masking. If, on the other hand layout is as fundamental to scene perception as gist and also contributes to conceptual masking, then changes in layout, irrespective of whether gist is (Experiment 2) or is not (Experiment 1) maintained, will automatically draw attention, causing conceptual masking.

METHOD

Observers saw a sequence of to-be-remembered target pictures (16 in Experiment 1 and 32 in Experiment 2) interspersed with the same number of to-be-ignored conceptual masks. In Experiment 1 ($N = 64$), similar to prior research, the conceptual masks presented a new gist each time. In one condition, the new objects always appeared in the same layout and in the other they appeared in novel layouts. In Experiment 2 ($N = 40$) the conceptual masks always presented the same gist (same objects on same background, e.g., bath products on counter), but again, in one condition the objects always appeared in the same layout (i.e., same conceptual mask repeated throughout the sequence) or they appeared in a new layout each time. Memory for target pictures was tested using a two-alternative forced choice test with similar distractors (two border collie dogs in a grass field) at the end of each sequence. To rule out any lower level masking differences between the pictures, across each experiment, the pictures

were counterbalanced such that each conceptual mask followed each target picture.

RESULTS

The percentage of correctly recognized target pictures for both experiments is shown in Figure 1. In Experiment 1, conceptual masks that changed in both gist and layout were harder to ignore, leading to impaired memory for target pictures than masks that changed only in gist, $t(62) = 3.10$, $p < .01$. In Experiment 2, when gist was always constant, conceptual masks that changed in layout were again harder to ignore, than when layout remained constant $t(38) = 3.16$, $p < .01$. Thus, the onset of a novel layout also captures attention, independent of gist. The results suggest that layout is rapidly and automatically extracted even when trying to ignore a scene. Repeating the same layout facilitated the observer's ability to reject the automatic processing elicited by a novel picture and continue processing the target picture.
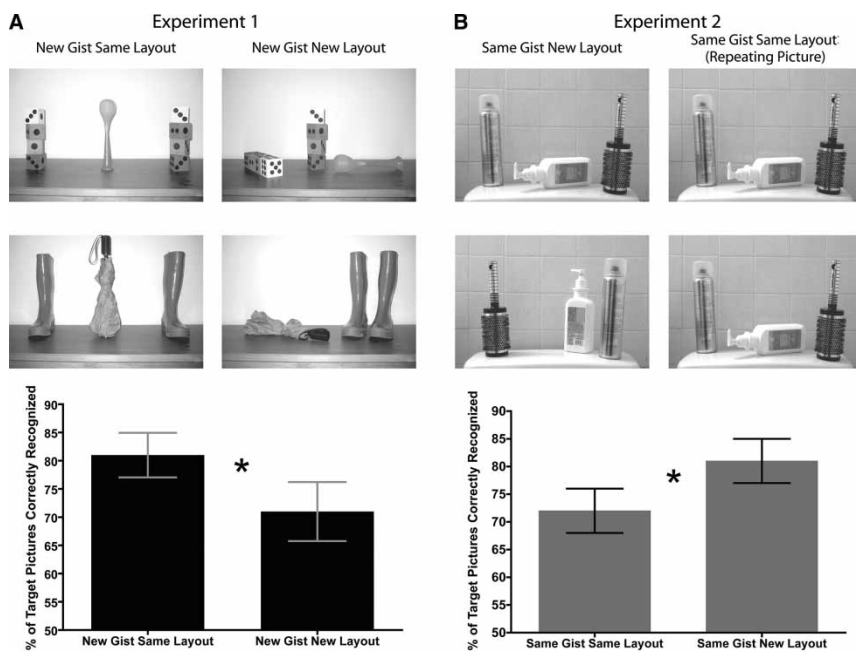

Figure 1.   Examples of the conceptual masks (top) and the percentage of correctly recognized target pictures (bottom) for each condition in Experiment 1 (A) and Experiment 2 (B). Error bars depict the 95% confidence interval for each mean.

DISCUSSION

Our results clearly demonstrate that the memory disruptions seen in conceptual masking are not solely due to the onset of a novel concept, as was previously argued. Rather, conceptual masking is also due to the onset of a novel layout. In Experiment 1, observers found it harder to ignore pictures that changed in both gist and layout than ones that changed only in gist. In Experiment 2, even when the same objects appeared on the same background, observers again found subtle changes in the layout of an otherwise repeated picture hard to ignore. These results demonstrate that layout is automatically grasped and influences the early processing stages of scene perception, even when observers actively ignore a scene.

In both Experiments 1 and 2, repetition of the same abstract layout made the conceptual masks easier to ignore. The implicit learning and use of layout information by our observers draws an interesting parallel between conceptual masking and contextual cueing. In contextual cueing, reaction times to find a target "T" embedded amongst "L" distractors become faster over time when the layout of the display is repeated (Chun & Jiang, 1998). In spite of observer's failure to explicitly recognize previously viewed displays, their visual system implicitly learns the layout and responds faster to repeated as compared to novel displays. Our results suggest that despite observer's attempts to ignore the conceptual masks, they automatically extracted layout information and formed an abstract representation of the layout that was independent of the objects. This abstract representation of layout might facilitate performance by allowing observers to more quickly and efficiently ignore the conceptual masks and continue processing the target pictures, thus leading to better consolidation and memory.

The study of scene perception has historically placed the emphasis on understanding gist. However, these results suggest that the abstract relations among objects in a scene are available rapidly and are automatically extracted. Our results demonstrate that layout plays an important and powerful role in our processing of scenes, even when they are being ignored.

REFERENCES

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.

Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 115–125.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 509–522.

A dual-processes model of attentional guidance for contextual cueing

Hirokazu Ogawa

Research Center for Advanced Science and Technology, University of Tokyo, Tokyo, Japan

Katsumi Watanabe

Research Center for Advanced Science and Technology, University of Tokyo, and ERATO, Japan Science and Technology Agency, Tokyo, Japan

Allocating visual attention to potentially important locations and objects in a visual scene is a critical function of the visual system. A substantial body of recent work indicates that the visual system utilizes visual memory to achieve efficient control of visual attention. One of the demonstrations showing that a memory-based process facilitates attentional acquisition of relevant information is known as contextual cueing (e.g., Chun & Jiang, 1998). Contextual cueing refers to faster reaction times to find a visual search target in displays where the locations of targets and distractors remain constant throughout the experiment, relative to when their locations are newly configured. Interestingly, the contextual cueing effect occurs without participants' awareness of the repetition of layouts. Contextual cueing appears to be due to an implicit process that learns invariant information in a visual scene and guides visual attention to the potential target locations.

In the present experiment, we examined how and when a target location is associated with a surrounding configuration of distractors in contextual cueing. There are many different processes involved in visual search. Which process is the most relevant for the association process of contextual cueing? One possibility is that a target location is associated with a representation of a display configuration formed during active visual search. Alternatively, a target location may be associated with spatial layouts at the moment when the target is detected. We examined these two possibilities in a contextual-cueing task. Immediately after participants made a saccade to a target, the searched layout was replaced by a different layout with the same target position. Thus, the layout in which participants had searched for the target

(searched layout) and the layout presented at the moment of target detection (nonsearched layout) differed.

METHODS

Participants were graduate students from the University of Tokyo. Eye movements were recorded with an Eyelink 2 tracker. The participant sat on a chair with the head stabilized by a headrest at 48 cm from the display. The stimuli were white "T" targets, distractors were "L"s, and premasks that were formed from the line segments of every possible target and distractor. They were drawn on a grey background. Targets consisted of $90°$ and $270°$ rotated "T"s. Distractors were normal or mirror-imaged "L"s rotated $0°$, $90°$, $180°$, or $270°$. Displays were generated by randomly placing items on an imaginary 20×20 grid. There were seven distractors and one target in each display.

Each trial was started by pressing a spacebar and began with fixating a white small circle at the centre of the display (Figure 1). At the beginning of the trial, a display from the searched display condition was presented, which consisted of eight premasks. When participants made a saccade and landed on within $1°$ distance from one of premasks, the premasks changed to a target or a distractor by removing line segments from the premask. When the eyes landed on the target, the whole display was masked for 100 ms and then the display was replaced with a display from the nonsearched display condition. The configuration of both searched and nonsearched displays was kept constant throughout the entire experiment. They shared the same target position but none of the distractor positions. The participant's task was to search the target by making saccades to premasks and to indicate the orientation of the target by pressing the appropriate keys.

The experiment consisted of a learning session (18 blocks), a test session (3 blocks), and a recognition session (3 blocks). At the beginning of the learning session, 32 repeated configurations were randomly generated and randomly assigned to the searched layouts condition or the nonsearched condition. In the learning session, the searched and nonsearched displays were paired and presented once in a block, resulting in 18 repetitions. Participants performed the visual search task only in the searched display. In the following test session, 16 searched, 16 nonsearched, and 16 newly generated displays were presented, and they performed the task for all displays. In the recognition session, participants were asked to report whether they recognized the layout and to guess target locations.
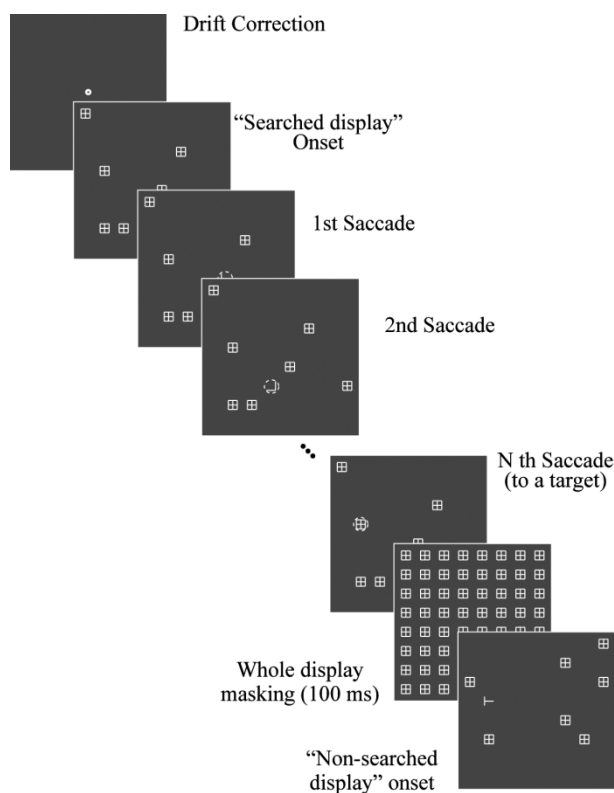
Figure 1. Sequence of events on a trial. Dotted circles indicate gaze points and were not presented in the actual display.

RESULT AND DISCUSSION

In the test session, the number of saccades to find targets was fewer in the searched displays than in the new displays, suggesting that the layouts were successfully learned during active search in the learning session and facilitated the visual search performance. Moreover, even fewer saccades were required to find targets in the nonsearched displays than in the new displays, indicating that contextual cues were also encoded at the moment of the target detection.

Although both searched and nonsearched layouts produced the contextual cueing effect, they affected eye movements during visual search. The contextual cue learned in the searched displays affected differently from the early stage of visual search (i.e., from the first saccade). Saccades tended to land closer to the target in the searched display from the first few saccades, and the landing locations became progressively closer to the target. On the other hand, the contextual cue learned in the nonsearched displays affected

the amplitude of the last saccade; this was larger in the nonsearched condition than in the other conditions. This pattern of eye movements indicates that the contextual information in the nonsearched displays became available only at the later stage of visual search.

Our results suggest that there may be two different processes of attentional guidance: The contextual information learned during active search is available from the beginning of visual search (i.e., from the first saccade) and gradually guides eye movements towards a target area. On the other hand, the contextual information learned at the moment of target detection becomes accessible later in the search process and directly indicates the target location.

REFERENCE

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28–71.

Enhanced redundant target effect in callosotomized individuals is not sensory in nature: Evidence from total and partial split-brain individuals

Catherine Ouimet and Pierre Jolicœur

Centre de Recherche en Neuropsychologie et Cognition, Université de Montréal, Montréal, QC, Canada

Jeff Miller

University of Otago, Otago, New Zealand

Alexia Ptito and Maryse Lassonde

Centre de Recherche en Neuropsychologie et Cognition, Université de Montréal, Montréal, QC, Canada

The aim of this study was to investigate the Redundant Target Effect (RTE) in callosotomized individuals. The RTE is a decrease of reaction times (RTs)

Address correspondence to Catherine Ouimet, Département de Psychologie, Université de Montréal, CP 6128, succursale Centre-ville, Montréal, Québec H3C 3J7, Canada. E-mail: catherine.ouimet.1@umontreal.ca

when two or more targets are shown rather than a single one. Two accounts for this phenomenon are the horse race and coactivation models. According to the horse race model, sensory information about each stimulus is carried through a separate channel (Raab, 1962). In the case of redundant targets, only the fastest of multiple stimuli is necessary to reach a criterion level of activation that triggers a motor response, where the rate of processing within a single channel is completely independent of the other one. Alternatively, the coactivation model assumes a neural mechanism summating activation from multiple channels (Miller, 1982).

Traditionally in the redundant condition of RTE studies with a visual simple RT task, one signal was presented to one hemisphere and the other signal was presented to the other hemisphere. Surprisingly, individuals whose corpus callosum had been surgically sectioned showed a greater RTE than callosally intact individuals (Reuter-Lorenz, Nozawa, Gazzaniga, & Hughes, 1995). From the neural coactivation perspective, such a finding is counter-intuitive, because the absence of corpus callosum in split-brain individuals should prevent efficient interhemispheric transfer and consequently reduce the likelihood of coactivation. The enhanced RTE found with bilateral presentation suggested that the activation of both hemispheres led to the occurrence of RTE (Miller, 2004).

In the present study, we tested individuals with complete sections of the corpus callosum and individuals with partial sections. Individuals with partial sections had intact posterior callosal pathways (i.e., the splenium was intact) and sectioned anterior fibers. Given that the posterior pathways connect sensory areas of the two hemispheres (e.g., Kolb & Whishaw, 1996), it was hypothesized that if the RTE occurs at a sensory level, partial split-brain individuals should show a profile of RTE similar to the one shown by normal individuals, because sensory information should transit normally through the intact posterior portion of their corpus callosum. In contrast, if the RTE occurs at a motor level, partial split-brain individuals should show a RTE profile more similar to the one presented by total split-brain individuals because they both lack the portion of the corpus callosum conveying motor information.

METHODS

Eight split-brain individuals were tested. Amongst them, four had under-gone complete callosotomy, and the other four had undergone anterior callosotomy. Ten neurologically intact individuals were also tested.

The stimuli were discs that occupied $2°$ of visual angle. Six possible locations were positioned around an imaginary circle about fixation to

Figure 1.   All the possible conditions as a function of the number of stimuli.

equate eccentricity as illustrated in Figure 1. Peripheral discs were located at
5° to the left or right of fixation.

The experiment was composed of six blocks. Two blocks had interhemi-
spheric conditions, two had intrahemispheric conditions, and two had a
midline condition. One interhemispheric block used discs presented in the
upper visual field, whereas the other used discs presented in the lower visual
field. One intrahemispheric block used discs presented in the right visual
field; the other used discs in the left visual field. In the midline condition,
stimuli were presented on the vertical meridian.

Each block was composed of 120 trials that were divided into four
equiprobable events (i.e., $-,-$), $(+,-)$, $(-,+)$, $(+,+)$, where "$+$"
means the signal appears and "$-$" means the signal does not appear. Trials
with eye movements were rejected.

Each participant was seated 40 cm in front of the computer screen. Each
trial began with a fixation point. It stayed on the screen for an average
interval of 500 ms, this duration varying randomly between 400 ms and 600
ms from trial to trial in order to avoid anticipation responses. Then, one or
two discs were presented for 200 ms while the fixation point was still
displayed on the screen. Participants were allowed a duration of 1000 ms to
press the response keys. As soon as a response was recorded, the fixation
cross disappeared and there was a 1000 ms interstimulus interval.

Participants responded bimanually by pressing the "Z" key with their left
index finger and the "M" key with their right index finger as soon as they

perceived any disc. This method allowed us to measure two responses per trial and to measure the RTs separately for each hand.

RESULTS

Analyses were run separately for split-brain individuals and normal individuals. The RTE was computed based on the fastest RT between the two hands. A significant RTE was found in split-brain individuals, $F(1, 6) = 29.99$, $MSE = 580.96$, $p < .002$, as well as in normal individuals, $F(1, 9) = 92.55$, $MSE = 66.97$, $p < .0001$. In split-brain individuals, the RTE was not modulated differently for partial and total split-brain individuals whether stimuli were presented on the midline or in intra- or interhemispheric fashion, $F(2, 12) = 0.05$, $MSE = 519.64$, $p > .95$, suggesting that the two groups of split-brain individuals were similar. Furthermore, a Condition (inter/intra/midline) × Number of stimuli (1 or 2) interaction was revealed, $F(2, 12) = 4.80$, $MSE = 519.64$, $p < .03$, suggesting that whether one or two hemispheres were directly activated by stimuli influenced the magnitude of the RTE. Specifically, the interhemispheric condition led to an enhanced RTE (38 ms) as compared to the intrahemispheric (10 ms) and midline (18 ms) conditions. It is interesting to note that such a difference was not found in normal individuals.

According to our results, the RTE pattern in partial split-brain individuals and total split-brain individuals is similar. Furthermore, despite the preservation of the posterior portion of their corpus callosum, partial split-brain individuals, just like total split-brain individuals, differ from the RTE pattern shown by callosally intact individuals. This pattern suggests that the absence of the anterior portion of the corpus callosum is crucial for enhanced RTE. As a consequence, it suggests that the RTE is likely to occur at a motor level, because the transfer, or the absence of transfer, of sensory information does not modulate the RTE pattern in split-brain individuals.

REFERENCES

Kolb, B., & Whishaw, I. Q. (1996). *Fundamentals of human neuropsychology* (4th ed.). New York: W. H. Freeman & Co.

Miller, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279.

Miller, J. O. (2004). Exaggerated redundancy gain in the split brain: A hemispheric coactivation account. *Cognitive Psychology*, *49*, 118–154.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*, 574–590.

Reuter-Lorenz, P. A., Nozawa, G., Gazzaniga, M. S., & Hughes, H. C. (1995). Fate of neglected targets: A chronometric analysis of redundant target effects in the bisected brain. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 211–230.

Effects of repetition on comprehending and remembering action pictures

Ann Rémond and Veronika Coltheart

Macquarie Centre for Cognitive Science, Macquarie University, Sydney, Australia

Differences between object and action processing have been documented in the psycholinguistic, developmental, and neuropsychological literature. These differences might be due to linguistic differences between objects and actions, for example, objects can be labelled using a specific noun whereas actions are described using a verb plus adverbs and verb particles to interpret the motion event. Tranel, Kemmerer, Adolphs, Damasio, and Damasio (2003) argued that when an action concept is presented, a pool of phonological or orthographic forms of appropriate words is activated. However, these action concepts may not exactly correspond to the meanings of verbs. Instead, the meanings of verbs many be "packaged" in language-specific ways so that interpretation is necessary to communicate action information. The generality of verb meanings makes mapping them onto a referent more difficult. Nouns, however, do not have this property, as only one word is often necessary to represent a single concrete entity. Gentner (1981) argues that verbs in general are more difficult to process than nouns because they have fewer components of meaning.

Neuropsychological studies of aphasic patients show that processing differences exist as double dissociations between the ability to recognize and name objects and actions (Breedin & Martin, 1996). However, most of what we know about picture memory comes from studies involving objects and scenes. The current research investigates whether significant processing differences occur between object and action pictures using a repetition blindness paradigm.

People are able to extract sufficient information to identify an object within only 100–200 ms of presentation (Potter & Levy, 1969). When stimuli are presented at rates of 10–16 items per second, comprehension is good while recall is poor unless this information is consolidated into longer term memory (Forster, 1970; Potter, 1976). Kanwisher (1987) observed that when

stimuli are presented in rapid serial visual presentation (RSVP) in series of three items where the first and third items are repeated, subjects are often only able to report one of the repeated items. She termed this failure to see or recall the second of the two identical items "repetition blindness" (RB). According to her Token Individuation hypothesis, type nodes are activated when the first instance of the stimulus is seen. However, successful report of the item requires a second stage to establish an episodic memory token and this process cannot occur for the item the second time within such a short interval, causing subjects to be unaware of, or "blind" to the repeated item.

The purpose of this study is to determine whether pictured actions are subject to repetition blindness or whether, because they are conceptually different from objects, they are not susceptible to the same cognitive processing limitations. A series of eight experiments were carried out using an RSVP paradigm typical in investigations of repetition blindness. Subjects viewed photographs of familiar actions obtained from Fiez and Tranel (1987) from a collection used to investigate action concept knowledge with aphasic patients. Photographs were shown at 114 ms/picture in three different conditions: 20 trials where the first and third picture were repeated, 20 trials where three different actions were shown, and 20 trials where only two actions were shown. Sequences were preceded and followed by masks and trial order was randomized.

In Experiments 1 and 2, subjects viewed action photographs in colour and greyscale, respectively. Results of both experiments indicated that subjects were significantly better at identifying both repeated actions, than at reporting both actions when the first and third actions were different. When subjects viewed the same photographs but were asked to remember the name of the object in the pictures (Experiments 3 and 4), the same repetition benefit was obtained. When all of the background information was removed from the picture (Experiments 5 and 6), subjects still found it easier to identify both repeated actions; however, when presented with objects (used in actions) alone, significant repetition blindness occurred (see Figure 1). Line drawings of actions (Experiment 7) produced a repetition benefit; however, with sequences of action words (Experiment 8), repetition blindness occurred.

Thus, a significant repetition benefit was found with pictures that depicted actions as static entities. This effect did not appear to be due to colour cues, contextual cues, or photographic quality of the stimuli. The results demonstrate processing differences between words and pictures because, when presented as written words, both nouns and verbs elicit repetition blindness and appear to be treated the same way. Pictured objects and actions are likely to be processed semantically prior to name retrieval, and actions appear to be immune to repetition deficits. Why might repeated actions be more readily perceived? Perhaps, because repetition is an integral
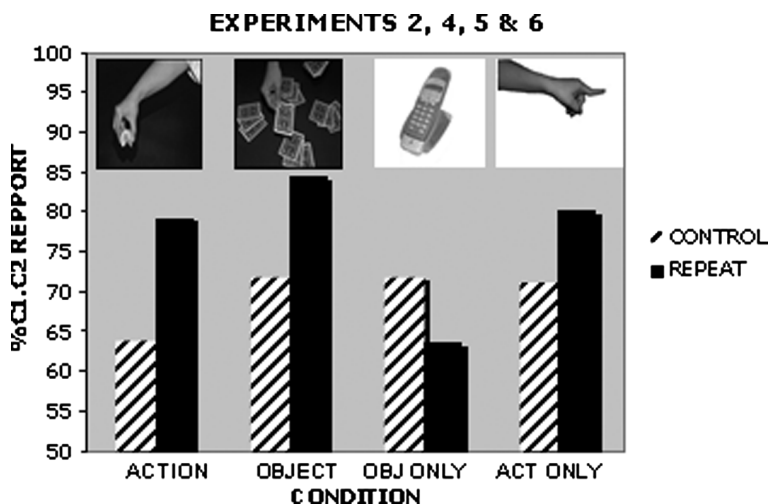
Figure 1. Summary of Experiment 2, showing a repetition benefit with action pictures; Experiment 4, showing a repetition benefit with object pictures; Experiment 5, showing a repetition blindness effect with object-only pictures; and Experiment 6, showing a repetition benefit with action-only pictures.

part of many actions, for example in locomotion, tool use, and speech, our visual cognitive systems have evolved special mechanisms to detect and perceive repetitions when these occur. Thus, for dynamic stimuli such as actions, the limits on token individuation obtained for static entities such as objects or written words may not arise. Investigation into the similarities and differences between object and action processing may further understanding of action concept knowledge and may contribute to theories of short-term memory, particularly early memory for rapidly presented information.

REFERENCES

Breedin, S. D., & Martin, R. C. (1996). Patterns of verb impairment in aphasia: An analysis of four cases. *Cognitive Neuropsychology, 13*, 51–91.

Fiez, J. A., & Tranel, D. (1987). Standardized stimuli and procedures for investigating the retrieval of lexical and conceptual knowledge for actions. *Memory and Cognition, 25*, 543–569.

Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception and Psychophysics, 8*(4), 215–221.

Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory, 4*, 161–178.

Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition, 27*, 117–143.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5), 509–522.

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

Tranel, D., Kemmerer, D., Adolphs, R., Damasio, H., & Damasio, A. R. (2003). Neural correlates of conceptual knowledge for actions. *Cognitive Neuropsychology*, 20(3/4/5/6), 409–432.

Subjects use configural information more than feature information to recognize inverted faces

Noah Z. Schwartz and Shuinn Chang

University of Southern California, Los Angeles, CA, USA

The face inversion effect (FIE) is a well-characterized phenomenon in which inverted faces are recognized slower and with less accuracy than upright faces (Yin, 1969). Numerous studies have supported the conclusion that upright faces are recognized in terms of the configural information intrinsic in the face whereas inverted faces are recognized in terms of the constituent features (Rhodes, Brake, & Atkinson, 1993; Tanaka & Farah, 1993). Though these results generally constitute the majority opinion, they are confounded by the fact that there is little control for task difficulty when subjects are discriminating between stimuli defined by feature differences versus configural differences. Rhodes et al. (1993), in fact, report a situation where varying task difficulty significantly changed the outcome of the experiment. Similarly, there is little control over individual differences in sensitivity to face information. This issue is of critical importance given the own-race effect in which subject performance depends on the familiarity of the race of the face being recognized (Tanaka, Kiefer, & Bukach, 2004).

Although it seems safe to assume that stimulus manipulations that appear obvious are equally salient to the observer, these results show that differences in task difficulty can influence the outcome of an experiment to a significant extent. In light of these complications, the original question remains as to how inverted face processing and recognition differs from that of upright faces.

Address correspondence to Noah Z. Schwartz, Department of Psychology, University of Southern California, SGM 501, Los Angeles, CA 90089-1061, USA. E-mail: nschwart@usc.edu

METHOD

This question was addressed using a novel memory paradigm known as Error-From-Sample (EFS) that uses highly calibrated stimuli to measure false alarm rates in subjects when performing a recognition task. EFS is a novel memory paradigm based on the traditional forced-choice (FC) paradigm in which subjects study a target and then select that target from a set of alternatives. In a normal FC trial, one of the alternatives is the target and the others are distractors that differ from the target in some controlled or uncontrolled fashion. In an EFS trial, however, the target is not presented as one of the alternatives yet the subject is led to believe that a correct response does in fact exist. Meanwhile, the distractors that are presented each vary from the target in a single stimulus dimension. These changes are calibrated to the sensitivity of the individual subject so that each distractor appears equally different from the target. As a result, the distractor that is chosen (the distractor that the subject believes to be the target) is an indication of which stimulus dimension is used least during the task. Over a number of trials, the distribution of errors committed across stimulus dimensions is a reflection of the underlying processing strategy being employed during the task. Because every EFS trial induces an error, the EFS design is known as a forced-error design or X-AFE.

The current study measured subject sensitivity to changes in nine internal face dimensions, and then applied the EFS method to determine which of those nine dimensions were being utilized to recognize upright and inverted faces. The dimensions used in the current study consisted of four feature dimensions and five configural dimensions. Feature dimensions were eyebrow shape, eye shape, nose shape, and mouth shape and were manipulated by morphing between two exemplars. Configural dimensions were eyebrow–eyebrow distance, eyebrow–eye distance, eye–eye distance, eye–nose distance, and nose–mouth distance and were manipulated by varying the relative position of each feature. Stimuli were extensively pretested over each morph and position range in order to ensure that discrimination thresholds were linear, equivalently interval in scale, and stable across the entire range of each dimension.

Discrimination thresholds were measured for the nine dimensions in both upright and inverted faces using a delayed match-to-sample same/different paradigm driven by the QUEST adaptive threshold procedure augmented by bias control trials. Trials were blocked by orientation and randomized across dimension. This method yielded a total of 18 thresholds that were used to calibrate distractors in a subsequent 3-AFE EFS task. Because EFS error frequency is inversely proportional to the utility of the dimension in which the errors were committed, one would expect subjects to commit few

configural errors and more feature errors when recognizing upright faces, while the opposite pattern is expected for inverted faces.

RESULTS AND DISCUSSION

When comparing grouped feature and configural dimensions, subjects committed significantly more feature errors than configural errors when recognizing upright faces (77.3% feature vs. 22.7% configural), $t(7) = 2.937$, $p = .022$. This result supports the hypothesis that upright faces are recognized in terms of their configural properties. Inverted faces, however, showed a similar trend where feature errors were committed more frequently than configural errors, suggesting that inverted faces are also recognized in terms of configural properties (68.5% feature vs. 31.5% configural), $t(7) = 2.528$, $p = .0393$. Despite the slight shift in error frequencies, there was no interaction between dimension type and stimulus orientation, $F(1, 14) = 0.302$, $p = .592$. This finding disagrees with existing theories that inverted faces are recognized in a feature-based manner.

When making orientation-related comparisons within each dimension, feature errors did decrease slightly as a function of face inversion but this shift only accounted for 4% of the difference in feature versus configural errors. In a 9×2 within-subjects analysis of variance, this shift was statistically significant, $F(8, 44) = 4.509$, $p = .0005$. Post hoc comparisons showed that all dimensions varied significantly with the exception of mouth shape. Error frequency per dimension is shown in Figure 1.
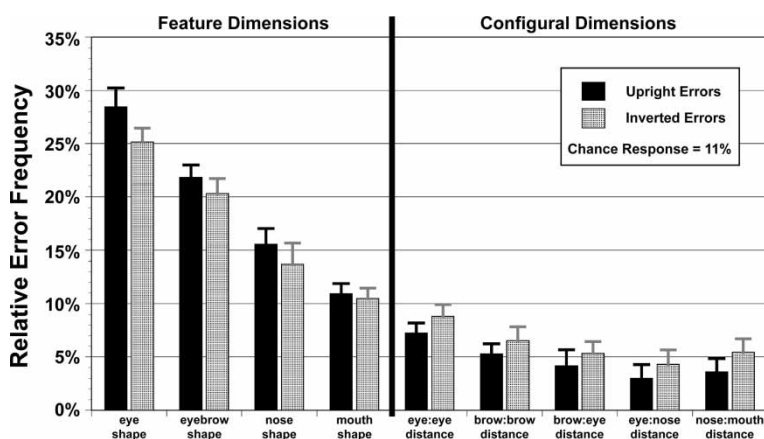

Figure 1. Relative frequency of errors committed in each face dimension as a function of orientation. Subjects commit more feature errors when recognizing faces regardless of orientation. Although there is a slight shift in error frequency for each individual dimension, grouped error frequencies did not vary significantly as a function of orientation.

It is also apparent, when viewing Figure 1, that feature errors did not cluster together in the same manner that configural errors did, possibly leading to the nonsignificant interaction described earlier. However, principle component analysis of overall error distribution yielded a single factor that loaded very high for configural dimensions and very low for feature dimensions, confirming the feature-configural dichotomy.

In summary, we find that after controlling for difficulty between conditions in a task that requires subjects to recognize faces that vary in either configural or feature dimensions, subjects appear to use configural information more than feature information for both upright and inverted faces.

REFERENCES

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*(1), 25–57.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*(2), 225–245.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*(1), B1–B9.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology: Human Perception and Performance*, 81, 141–145.

Learning to form new perceptual groups

Timothy J. Vickery and Joshua K. Hartshorne

Department of Psychology, Harvard University, Cambridge, MA, USA

Yuhong V. Jiang

Department of Psychology, University of Minnesota, Minneapolis, MN, USA

Demonstrations such as R. C. James' well-known black-and-white Dalmatian image suggest that experienced structure plays a central role in

Address correspondence to Timothy J. Vickery, Vision Sciences Lab., Department of Psychology, Harvard University, 33 Kirkland St., 7th floor, Cambridge, MA 02138, USA. E-mail: tim.vickery@gmail.com

perceptual grouping. At first glance, the scene in this image is difficult to recognize. However, if told what the scene contains, observers can easily see a Dalmatian standing on a leafy lawn. Knowledge of the shape of a dog enables grouping of elements that are otherwise too similar and proximal to irrelevant portions of the image to lead to recognition. Although some prior studies have examined the effects of learning in grouping (e.g., Zemel, Behrmann, Mozer, & Bavelier, 2002), the processes that underlie such experiential effects are still poorly understood. For example, although it is known that humans can recognize statistically cooccurring shape pairs (Fiser & Aslin, 2001), it remains unknown whether the statistical cooccurrence of two shapes can induce grouping of those shapes. In this study, we showed that when shape pairs cooccurred repeatedly and consistently along with a common region grouping cue (Palmer, 1992) that always grouped them together, residual grouping could later be measured in the absence of an explicit grouping cue.

To test the hypothesis that the pairing of shapes with external cues can induce long-term grouping effects, we turned to an objective method of studying grouping introduced by Palmer and Beck (2007). In this task, a row of elements is presented (e.g., squares and circles), which are grouped by some cue (e.g., shape) into pairs. The elements alternate on some other property (e.g., colour) except for once, when the property is repeated (e.g., two adjacent white elements). The goal is to locate or identify this repetition. Participants are slower at detecting such repetitions when they cross group boundaries than when they fall within group boundaries. In the example used above, if the white-white repetition crosses a shape boundary (a square and a circle are white), participants are slower to detect this repetition than if the repetition occurs across two shapes that are the same. This result generalizes (at a minimum) to grouping by similarity, proximity, common region, and connectedness.

In our primary experiment, we exposed participants to a training phase in order to induce grouping, followed by a test phase to assess grouping after the removal of external cues (Figure 1 depicts trials from both phases). Our training phase consisted of exposure to displays composed of 24 unique novel shapes (borrowed from Fiser & Aslin, 2001). For each participant, the 24 shapes were randomly divided into six sets of 4 shapes. Each trial contained only the 4 shapes from one particular set (e.g., shapes A, B, C, and D) repeated in order with the first shape chosen at random, such that each scene contained a row of 15 shapes. These shapes were grouped, pairwise, by a common region grouping cue (a black frame rectangle that surrounded both shapes) such that shape pairs were consistently grouped (e.g., A and B were always grouped, as were C and D). The shapes alternated in colour (light grey and white), with a single repetition of colour in every scene, and the participants task was simply to press a key when it was located, then
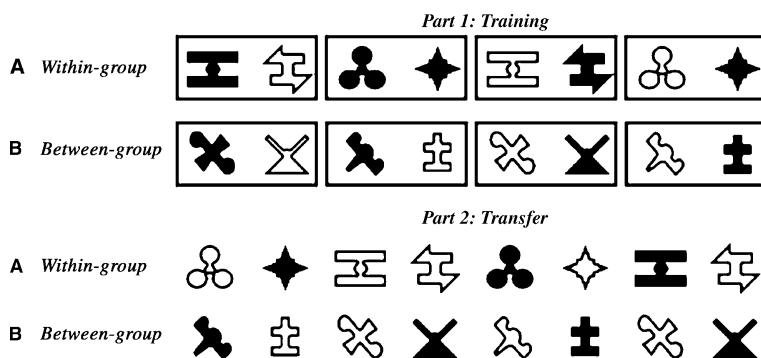
Figure 1.   Depictions of trials from each of the two conditions and two phases of the experiment. All four shapes in each of six sets occurred together on every trial in which any of them appeared. The common region cue consistently paired the shapes during training, and did not appear during transfer. Shapes are depicted in different number and colours than in the actual experiments (see text).

click with the mouse on placeholders that appeared in that position. The colour repetition was equally likely to occur within group boundaries (e.g., on shapes A and B, or C and D) as between group boundaries (e.g., B and C, or D and A). Participants were exposed 120 trials in each condition, divided equally amongst the six possible stimulus sets.

The training phase was followed by a transfer phase, which differed from training in three ways. First, the common region cue to grouping was removed from every display. Thus, any grouping effect observed was due to either shape similarity (which should be counterbalanced across subjects, since shape pairs were randomly assigned), or the effect of experiencing these shapes as grouped together. Second, the number of items per display was reduced to 11, so that variance in RT would be reduced. Finally, the number of total trials per condition was reduced to 60.

Results of the training phase showed that our grouping manipulation was successful in producing a grouping effect on reaction time (RT): Within-group RTs were substantially faster than between-group RTs due to the common region cue, with average RTs of 2114 ms and 2734 ms, respectively, $t(6) = 4.37$, $p < .01$. Results of the transfer phase showed that reaction time was faster when the colour repetition was within a group boundary than when it crossed a group boundary as defined by the common region grouping cue employed during training (1289 ms and 1422 ms, respectively), $t(6) = 2.94$, $p < .05$.

A follow-up experiment replicated these results and showed that this learning is highly specific. We constructed a set of stimuli (similar to those of the first experiment) that were symmetric across the vertical axis, but asymmetric across the horizontal axis. One group ($N = 15$) was trained and

tested with these stimuli just as in the above experiment. Their results showed similar effects of learning, with a residual effect of the grouping cue ($p < .05$). A second group ($N = 20$) was trained with the same set of stimuli, and tested in the transfer phase with vertically flipped elements. These participants failed to show a significant effect of learning (between-groups interaction, $p < .05$), suggesting that the learning effect is highly specific to the exact shapes that produce it.

These results demonstrate a novel example of learned grouping. If two shapes cooccur along with a grouping cue, then grouping effects on performance can be measured. If participants are repeatedly exposed to these scenes, this grouping effect can persist even when the explicit cue to grouping is removed from the display. This finding is an important step towards understanding how experience modifies perceptual organization, which in turn shapes the manner in which attention is deployed.

REFERENCES

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*, 499–504.

Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology*, *24*, 436–447.

Palmer, S. E., & Beck, D. M. (2007). The repetition discrimination task: An objective method for studying perceptual grouping. *Perception and Psychophysics*, *69*(1), 68–78.

Zemel, R., Behrmann, M., Mozer, M. C., & Bavelier, D. (2002). Experience-dependent perceptual grouping and object-based attention. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 202–217.

Integrating object locations in the memory representation of a spatial layout

Naohide Yamamoto and Amy L. Shelton

Johns Hopkins University, Baltimore, MD, USA

Spatial learning in everyday environments involves remembering the layout of multiple objects. Because few environments are learned at a single

Address correspondence to Naohide Yamamoto, Department of Psychology, George Washington University, 2125 G Street NW, Washington, DC 20052, USA. E-mail: nyama@gwu.edu

fixation, visual learning of a spatial layout requires integrating sequentially learned object locations into a representation of the entire layout (e.g., Henderson & Hollingworth, 1998). The present study investigated how this integration occurs by examining differential effects of two spatial learning methods on subsequent memory representations of spatial layout.

Previously we have demonstrated that a spatial layout can be learned efficiently through sequential viewing of objects (Yamamoto & Shelton, 2007, in press). In these studies, stationary observers were presented with six objects (each in unique location) sequentially, one or two at a time. When they subsequently made judgements of relative direction among objects (JRDs), their performance was equivalent or even superior to that following simultaneous viewing of the entire layout. Because this task primarily requires knowledge of interobject spatial relations, these findings suggest that there was little cost in merging sequentially experienced object locations into a representation of the spatial layout.

However, previous studies have shown that object locations in a given space are often organized into several collective units in memory that are not easily integrated (e.g., Hirtle & Jonides, 1985; McNamara, 1986; Wang & Brockmole, 2003). For example, Wang and Brockmole (2003) demonstrated that judgements of egocentric direction were less accurate for familiar locations on a college campus than for object locations in the immediate surroundings (i.e., a room in a building on the campus), even after both types of target locations were learned to the same criterion. These results suggest that not all locations can be integrated into a single spatial representation, even when all of them could be located within the same spatial framework.

To resolve this discrepancy, two experiments were conducted in the present study. In Experiment 1, stationary participants (four males and four females) were presented with a room-sized layout of 10 objects, which was divided into two sets of five objects. These two sets did not share any objects or locations, but they were also not linearly separable within the larger room context. The participants were instructed to remember the layout of all 10 objects, not two separate layouts of five objects each. They were shown the first set for 30 s, and asked to point to and name the five objects with their eyes closed. They repeated this study-test sequence until they fluently pointed to correct object locations twice in a row. Then the just-viewed set was removed and the same procedure was repeated for the second set.

After the learning phase, the participants performed JRDs. Three objects in the learned layout formed each trial, e.g., "Imagine you are at the bag and facing the jar. Point to the vase." The first two objects constituted an imagined heading; the third was a target. The major dependent variable was absolute angular error in pointing. The primary independent variable was the type of JRD trials: (1) All objects constituting a trial were from the same set (labelled as AA-A; for details of the trial labels, see the caption of
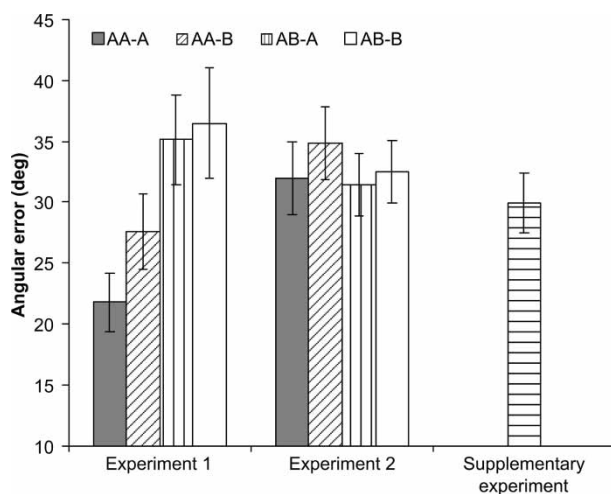
Figure 1.   Mean absolute angular errors in judgements of relative direction (JRDs) as a function of JRD trial types. In AA-A trials, all three objects constituting a trial came from the same set. The first, second, and third letters in the label represent a base object, a facing object, and a target object in a JRD trial, respectively. The same identity shared by all three letters indicates that those objects belonged to the same set. In AA-B trials, an imagined heading was defined within each set and a target object was from the other set. In AB-A and AB-B trials, the imagined heading was made up with objects from both sets (one from each), and the target object belonged to the same set as either the base object or the facing object. Note that in these trial labels, A and B simply indicate whether each object came from the same or different set, and they do not refer to objects from specific sets. In the supplementary experiment no distinction was made in the JRD trial types (i.e., all 10 objects were presented simultaneously), and therefore the overall mean of all JRD trials is plotted. Error bars represent ± 1 standard errors of the means.

Figure 1); (2) an imagined heading was defined by objects from the same set, and a target was from the other set (AA-B); and (3 and 4) an imagined heading was made up with objects from both sets (AB-A or AB-B). These types of JRD trials required the integration of two sets into a single layout to different degrees: AA-A trials could be performed without integration; AA-B trials required the integration only to find the target; and AB-A/AB-B trials necessitated the integration both for establishing the imagined heading and for locating the target. Therefore, by comparing performance in these different trial types, the present experiments explored how effectively two groups of object locations were integrated in memory. The JRD trials were presented in random order, and different imagined headings and target directions were counterbalanced.

Results showed that AA-A trials yielded most accurate JRDs, followed by AA-B and AB-A/AB-B trials in this order (see Figure 1), $F(3, 18) = 5.86$, $p < .02$. These results correspond to the differential amounts of required integration for each trial type, indicating that there was clear cost in

integrating two sets into a single representation of the entire layout. That is, although participants were explicitly instructed to put two sets together into a single layout prior to the learning phase and, in addition, they viewed all objects in the same room from the fixed viewing position, such integration still required additional cognitive processes when each set was learned individually.

In Experiment 2, the same procedure and instructions as in Experiment 1 were used with the following modification: After viewing the first set once, it was removed and the second set was presented immediately, and then participants (four males and four females) pointed and named all 10 objects with their eyes closed. (In contrast, participants in Experiment 1 learned the first set to criterion, and then learned the second set to criterion.) By presenting two sets in succession and setting the learning criterion for the whole layout, this procedure was intended to facilitate the integration of two sets during the learning phase. Results showed that although overall accuracy was decreased, all trial types yielded equivalent performance (see Figure 1), $F < 1$. In addition, these JRDs were as accurate as those performed after viewing all 10 objects simultaneously (supplementary experiment; see Figure 1). Together, these results indicate that the integration of two sets was done with little cost in Experiment 2.

Because the key difference between Experiments 1 and 2 was whether two sets were learned individually or successively, the present findings suggest that integration of object locations into the representation of a single layout can be performed effectively if it is carried out during initial encoding of the environment. However, the same integration requires additional processes if it is attempted at the time of retrieval, suggesting that separately formed spatial representations remain independent in memory, even when those representations have a large overlap.

REFERENCES

Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269–293). Oxford, UK: Elsevier.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory and Cognition*, *13*, 208–217.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*, 87–121.

Wang, R. F., & Brockmole, J. R. (2003). Simultaneous spatial updating in nested environments. *Psychonomic Bulletin and Review*, *10*, 981–986.

Yamamoto, N., & Shelton, A. L. (2007). Path information effects in visual and proprioceptive spatial learning. *Acta Psychologica*, *125*, 346–360.

Yamamoto, N., & Shelton, A. L. (in press). Sequential versus simultaneous viewing of an environment: Effects of focal attention to individual object locations on visual spatial learning. *Visual Cognition*.