Object Perception, Attention, and Memory (OPAM) 2006 Conference Report

14th Annual Meeting, Houston, TX, USA

Organized by Kate Arrington, Monica Castelhano, Steve Franconeri, and Andrew Leber

The role of global layout in visual short-term memory

George Alvarez and Aude Oliva

Massachusetts Institute of Technology, Cambridge, MA, USA

Among the most fundamental questions we can address in visual cognition research is what the basic unit of analysis is for a particular visual function. Work on visual memory has contrasted three candidate units: objects, spatial locations, and feature dimensions. An implicit assumption in this search for the basic units of memory is that the units are independent: whether objects, locations, or features are the basic unit of analysis, presumably each unit is encoded and stored in memory independently of the others.

Although some earlier research suggested that visual short-term memory operates over discrete objects (Lee & Chun, 2001; Luck & Vogel, 1997), more recent work suggests that the independence assumption may be false. Specifically, it appears that the spatial relationship between objects plays a role in memory for both spatial and featural information about objects. For example, disrupting the spatial layout of a display between the study display and the test display interferes with retrieval of both the location and the identity of objects (e.g., Jiang, Olson, & Chun, 2000).

Please address correspondence to George A. Alvarez, Room 4078c, Building 46, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: alvarez@mit.edu

While it appears that spatial configuration may play an important role in visual memory, it is less clear how we should characterize the spatial configuration of a collection of objects. Exactly what constitutes the spatial configuration of a collection of objects, and which dimensions of this configuration are important for memory? In the present study we attempt to quantify the spatial layout of objects and to determine which aspects of spatial layout are represented in memory.

QUANTIFYING SPATIAL LAYOUT

The first step for this study was to develop a formal method to quantify the spatial relationship between objects with a psychologically relevant measure. Gestalt psychology suggests many candidate features, and we began with what we call "spatial regularity", which represents the degree of regularity in the spacing between objects.

To quantify spatial regularity, we begin by reducing each object to a set of {x,y} coordinates corresponding to the object's centre of gravity. An index of spatial variability was computed by measuring the distance between each pair of objects, then taking the standard deviation of these distances and dividing it by the mean (variability index = std/mean). Dividing by the mean makes this measure a scale invariant index of the variability in the spacing between objects. Finally, an index of *spatial regularity* was taken as 1/(variability index). Displays with higher values on this regularity index tend to look more organized and structured (see Figure 1a).
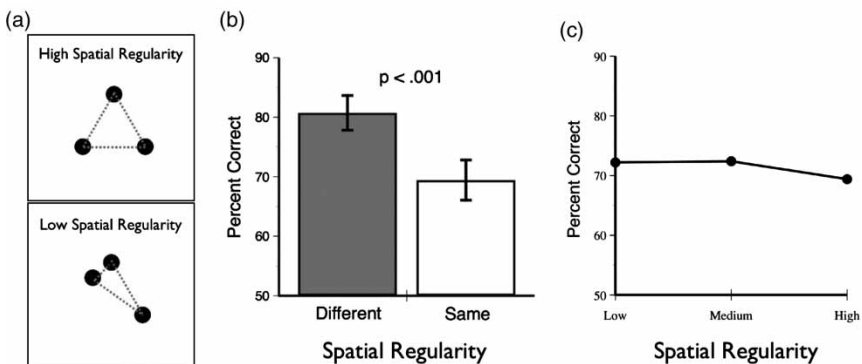

Figure 1. (a) Quantifying spatial regularity. (b) Percent correct on the change detection task was higher when the change altered spatial regularity (different) than when it left the spatial regularity unchanged (same). (c) Overall level of spatial regularity did not affect change detection accuracy.

REPRESENTATION OF SPATIAL REGULARITY IN MEMORY

We investigated the extent to which our measure of spatial regularity captures the representation of the spatial relationship between objects in memory. We had observers remember the location of a set of briefly presented objects, followed by a blank interval, and then a test display in which one item changed location. Two items in the test display were highlighted (the one that changed, and one that did not), and the task was to indicate which of the two items had changed location.

The distance of the location change was always 4 degrees visual angle. However, in the spatial regularity–same condition, the change resulted in the same spatial regularity between items, whereas in the spatial regularity–different condition the change altered the spatial regularity between items. The results showed that a location change that disrupts the spatial regularity was identified more accurately than one that did not alter the spatial regularity (Figure 1b). Critically, the local change in terms of pixel distance was identical for these two conditions, so it must be the magnitude of change in the spatial layout that distinguished between them.

It appears that the spatial relationship between objects is an important component of the memory representation, and our spatial regularity measure captures an important component of this representation. While previous work has suggested that spatial layout is encoded in memory (Jiang et al., 2000), to our knowledge the spatial regularity measure is the first formal representation that has been shown to capture a psychologically relevant component of spatial layout.

EFFECTS OF SPATIAL REGULARITY ON MEMORY CAPACITY

Although it appears that spatial regularity measures a critical component of relational encoding in visual memory, it does not appear to be a factor in determining the amount of information that can be stored in short term memory. Our results suggest that is just as easy to remember the layout of objects with low spatial regularity, as it is to remember the layout of objects with high spatial regularity (Figure 1c). Thus, while these regularities are part of the memory representation (as described above), they do not impact the fidelity of memory storage.

This indicates that the spatial relationships between objects are not encoded for the purposes of "saving memory space" by forming a more compact code for memory storage. Instead, these spatial relationships must play a different role. One possibility is that the degree of spatial regularity is an informative feature in real-world contexts. If this were the case, then the ability to detect and maintain this information in memory could be

important for learning and could lead to faster more efficient processing of information in real world displays. Further research exploring spatial regularity in real-world scenes is necessary to test this possibility.

REFERENCES

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 638–702.

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception and Psychophysics*, *63*(2), 253–257.

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

An information processing investigation of hierarchical form perception: Evidence for parallel processing

Leslie M. Blaha

Indiana University, Bloomington, IN, USA

Shannon A. Johnson

Dalhousie University, Halifax, NC, Canada

James T. Townsend

Indiana University, Bloomington, IN, USA

Studies of hierarchical form perception (e.g., large letter composed of individual letters) have explored many experimental conditions supporting findings of global advantage, i.e., global reaction times (RTs) are faster than local RTs, and global interference, where incongruent global information slows local processing; such studies have pursued general underlying processing mechanisms for these findings (for review, see Kimchi, 1992). There exists a powerful set of models and methodologies developed by Townsend and colleagues (Townsend, 1974; Townsend & Schweickert, 1989) designed to establish more detailed mechanisms involved in cognitive processing and perception.

Please address correspondence to Leslie M. Blaha, Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th Street, Bloomington, Indiana 47405, USA. E-mail: lblaha@indiana.edu

Our human information processing perspective identifies four fundamental characteristics of information processing systems: (1) architecture, (2) stopping rule, (3) (in)dependence, and (4) capacity (Townsend, 1974). Processing architecture is the spatial and temporal organization of subprocesses, the two major classes being sequential (serial) or simultaneous (parallel) structures. The stopping rule defines how much information is processed before a response is made. Processing dependency refers to potential interactions between subprocesses. Capacity refers to the workload efficiency of the system.

This approach enables the definition of basic cognitive mechanisms giving rise to observed global–local perception phenomena. We begin with a noninterference, divided attention paradigm to establish processing architecture, stopping rule and capacity under experimental conditions allowing for maximal facilitation (via information redundancy) between subprocesses. Analyses, then, focus on response time data for global-only, local-only, and redundant information conditions, to determine if RT advantages result from, for example, global information processing first in a serial system or by the global subprocess finishing first within a parallel system.

METHOD AND MEASURES

We employed systems factorial technology (SFT) methods developed by Townsend and colleagues (Townsend & Nozawa, 1995; Townsend & Wenger, 2004). Specifically, we applied the double factorial experimental paradigm (DFP), which is based on the factorial manipulation of two stimulus characteristics within a divided attention task. Shown in Figure 1, the hierarchical images were large arrows composed of smaller arrow elements. The presence or absence of a right-pointing arrow at either the global or local level (factor 1) provided a work-load manipulation necessary for measuring capacity (characteristic 4). Factor 2 was the salience of the arrow relative to the dash (low salience being very similar in appearance to the dash and high salience arrows being more distinct). Salience manipulations influence processing speed, which allowed us to assess both architecture (characteristic 1) and stopping rule (2). Participants were asked to respond "yes" if any right-pointing arrow was present in the stimulus, and "no" if no arrows were present (the dash composed of dashes).

Architecture and stopping rule were measured by both a mean interaction contrast (MIC) and survivor interaction contrast (SIC), which are double differences in the RT means or survivor functions, respectively, of the four factorial conditions comprised of redundant target stimuli (global right arrows composed of local right arrows). Each theoretical architecture and stopping rule combination makes a unique prediction for the MIC and SIC
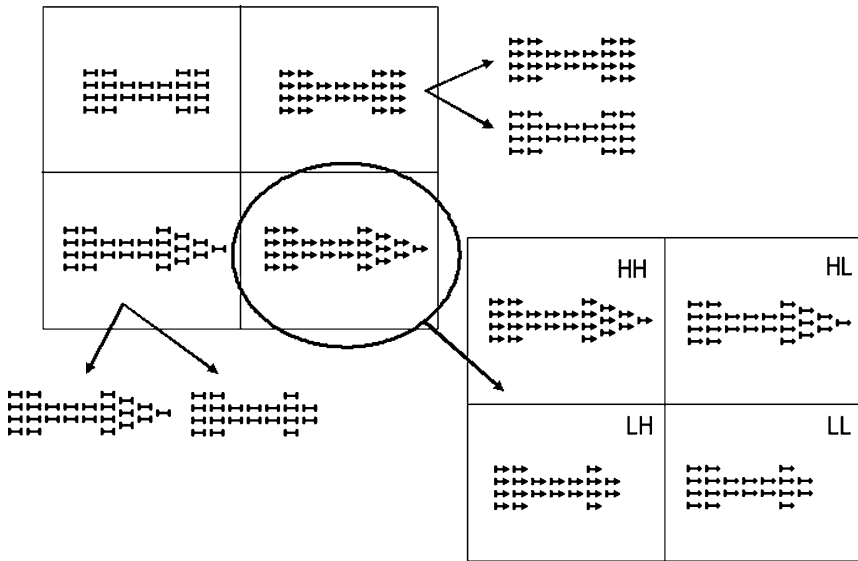
Figure 1. Schematic depicting the double factorial paradigm. Factor 1 was the presence/absence of a right-pointing arrow at the local and global levels (upper box). Factor 2 was a salience manipulation on the arrows. Factorial combinations of both arrow presence and arrow saliency form the critical conditions for SFT measures.

(Townsend & Nozawa, 1995), thus allowing us to compare observed measures with model predictions. Processing capacity was measured by the capacity coefficient (Wenger & Townsend, 2000), which is a ratio of processing in the redundant target stimuli to the sum of individual target images. That is, we take a log transformation of the RT distribution in the redundant arrows condition and compare it to the sum of log transformed distributions for the global-only and local-only conditions.

RESULTS

For the establishment of information processing systems, analyses focus on individual data. Thus, we report here trends observed within the data for our group of 14 adult participants. Reported trends are significant at the $p < .05$ level.

For 10 of 14 participants, mean RTs for the global-only condition were significantly faster than the local-only condition; 3 individuals exhibited faster local-only mean RTs. One participant had equal global-only and local-only mean RTs. Additionally, all participants exhibited a redundancy gain,

where mean RTs to the redundant arrows condition were faster than both the local-only and global-only conditions.

SFT measures were applied at the individual level; eight of our fourteen participants had significant effects for both experimental factors (arrow presence, salience), meeting model assumptions necessary for SFT measures.

All eight participants exhibited a parallel processing architecture. Six of these participants exhibited RT patterns consistent with a self-terminating stopping rule, meaning they responded as soon as either the global or local subprocess completed. Among these six participants, four individuals processed in an unlimited capacity manner, sometimes showing mild capacity limitations. One individual exhibited some early RT super capacity processing, with unlimited or mildly limited capacity for longer RTs. The sixth individual exhibited super capacity over most RT values; this super capacity forced RT distribution violations of bounds on unlimited capacity processing, signifying extreme super capacity processing. Thus, in general, participants simultaneously processed global and local information and terminated processing when minimal information had been accumulated to respond. As information increased in the stimuli, processing speed was either mildly slowed or not affected, with the exception of two individuals who were helped by increasing information. Note that these modelling results alone do not confirm which subprocess finished first (global or local).

A small subset of participants (two out of eight) exhibited parallel processing with an exhaustive stopping rule, meaning they processed both global and local information before responding. These two individuals also exhibited super capacity processing, with RT distributions that also violated the upper bounds on unlimited capacity processing. These individuals were, thus, extremely efficient at handling increases in the cognitive work load.

DISCUSSION

Hierarchical form perception usually involves parallel, self-terminating information processing with mild or no limitations in processing efficiency. Although we did find some heterogeneity among participants, this appears to be the dominant processing strategy in the absence of interfering stimulus information. These findings provide a parallel processing foundation for a complete model of hierarchical form perception. Selective attention tasks (so-called Garner filtering tasks), will examine potential subprocess inter-actions and processing under varying work-load demands, to further our understanding of both global advantage and global interference (see also Johnson et al., 2005).

REFERENCES

Johnson, S. A., Blaha, L. M., Hernandez-Ritter, I., Primeau, M., Murphy, R. R., Townsend, J. T., & Stout, J. C. (2005, May). *New perspectives on global-local processing in autism spectrum disorders*. Poster presented at the international meeting for Autism Research, Boston.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, *112*(1), 24–38.

Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 133–168). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*(4), 321–359.

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, *33*, 309–327.

Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, *111*(4), 1003–1035.

Wenger, M. J., & Townsend, J. T. (2000). Basic response time tools for studying general processing capacity in attention, perception, and cognition. *Journal of General Psychology*, *127*(1), 67–99.

Early crossmodal multitasking interference revealed by event-related potentials

Benoit Brisson and Pierre Jolicoeur

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

In the psychological refractory period (PRP) paradigm, two distinct targets, often presented in different sensory modalities, are separated by various time intervals (i.e., stimulus onset asynchronies, or SOA), and a speeded response is required for each target. Even with very simple stimuli and associated tasks, the PRP paradigm yields robust interference effects, reflected mostly by an increase in mean response time to the second target as SOA is reduced (i.e., as task overlap increases).

Virtually all models of dual-task interference agree that this lengthening of the response time to the second target is caused by interference at a late,

Please address correspondence to Benoit Brisson, Département de Psychologie, Université de Montréal, CP 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7, Canada. E-mail: benoit.brisson@umontreal.ca

central locus of processing, and most likely involves response selection (e.g., Pashler, 1994). It has also been argued that the central attention mechanisms that underlie the PRP effect are distinct, and therefore independent, of the mechanisms involved in deploying visual-spatial attention (e.g., Pashler, 1991). To investigate these claims, we measured event-related potential (ERP) responses to a lateralized visual target while observers performed a PRP paradigm. We focused on two ERP components: the N2pc (N2 *posterior contralateral*) and the sustained posterior contralateral negativity (SPCN). These two ERP components are thought to index visual activity for two main reasons. First, they arise at electrode sites contralateral to the to-be-processed visual items, which link the activity to the location of the task relevant items in the visual field. Second, they have a posterior scalp distribution, which is consistent with activity in extrastriate visual cortex. It is widely accepted that the N2pc is a valid index of the moment-by-moment deployment of covert visual-spatial attention (e.g., Eimer, 1996; Luck & Hillyard, 1994), while the SPCN, which follows the N2pc, is thought to reflect visual short-term memory activity (e.g., Jolicoeur, Brisson, & Robitaille, 2006; Vogel & Machizawa, 2004).

In the present experiment, a tone was presented in each trial, followed by a visual display that consisted of four squares, two on each side of fixation. Each square was red or green and had a gap in one of its sides. The first target was the tone and the second target (which appeared randomly in each of the four possible positions) was a red square amongst green distractors for half of the participants and a green square amongst red distractors for the other half. The first task was a speeded four-alternative discrimination of the pitch of the tone. The second task was a speeded four-alternative discrimination of the location of the gap in the visual target (up, down, left, or right). The tone-visual display SOA was 300 ms, 650 ms, or 1000 ms.

As is typically observed in PRP studies, there was no SOA effect on report accuracy of the second target. However, mean response time to the second target was lengthened as SOA was reduced [984 ms, 763 ms, and 717 ms for the 300 ms SOA, 650 ms SOA, and 1000 ms SOA, respectively; $F(2, 14) = 75.6$, $p < .0001$]. The observed PRP effect suggests that central processing of the tone interfered with some aspects of second task processing. To assess more precisely where the interference started to occur, we focused on the N2pc and SPCN.

To isolate the N2pc and the SPCN from overlapping activity that was not lateralized with respect to the side of the second target, the N2pc and SPCN were quantified following the subtraction of the ipsilateral waveforms from the contralateral waveforms (see Figure 1). A progressive attenuation of N2pc mean amplitude (in the 180–260 ms postvisual display time window) was observed as SOA was reduced, $F(2, 14) = 6.46$, $p < .03$, providing strong evidence that concurrent processing of the tone impaired visual-spatial
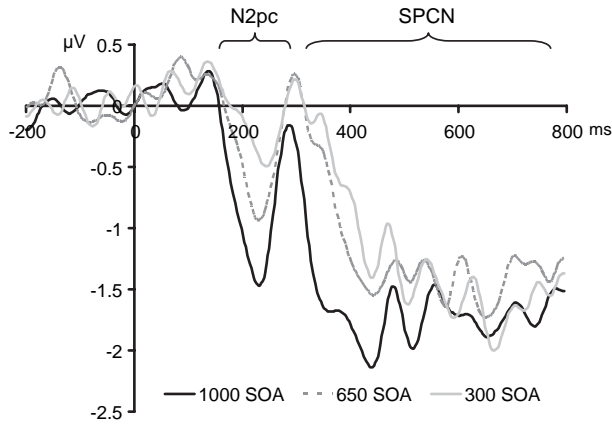
Figure 1.   Contralateral minus ipsilateral difference waves. Presented is the pooling of O1/O2, PO7/PO8, and P7/P8 electrode sites. A 15 Hz low-pass filter was applied for illustrative purposes.

attention. Furthermore, the attenuation of the N2pc was still significant when the longest SOA was removed, $F(1, 7) = 11.9$, $p < .01$. A differential pretrial preparatory state between conditions is not possible because SOAs were randomly intermixed in each block of trials. Furthermore, the second target onset occurred before mean response time to the first target in the two shortest SOA conditions and only 118 ms after mean response time to the first target in the longest SOA condition, which should leave too little time to increase second task preparation. Central processing, postulated to be responsible for the behavioural PRP effect, is therefore most likely also responsible for the observed progressive attenuation of the N2pc.

Another important discovery is that the SPCN onset latency was progressively delayed as SOA decreased. This observation, which was statistically confirmed by a jackknife analysis, $F(2, 14) = 10.5$, $p < .002$, strongly suggests that transfer into visual short-term memory was delayed. It is likely that the PRP effect is dominated by central postponement, because the increase in SPCN latency accounts for about a third of the behavioural PRP effect. Nevertheless, the substantial increase in the SPCN latency observed here is the first demonstration that the transfer of information into visual short-term memory can be delayed by a concurrent demanding speeded auditory task. Although delayed, the null effect of SOA on the amplitude of the SPCN in the later 550–750 ms postvisual display time window shows that a stable visual-short term memory representation was eventually achieved in all conditions, which was corroborated by the equivalent accuracy rates across SOA. In summary, the present results show that, contrary to what is widely assumed by prominent models of dual-task interference, the deployment of visual-spatial attention is impaired, and

encoding in visual short-term memory is delayed quite significantly, by crossmodal multitasking in the PRP paradigm.

REFERENCES

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*, 225–234.

Jolicoeur, P., Brisson, B., & Robitaille, N. (2006). *Dissociation of the N2pc and sustained posterior contralateral negativity in a speeded choice response task*. Manuscript submitted for publication.

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

Pashler, H. (1991). Shifting visual spatial attention and selecting motor responses: Distinct attentional mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 1023–1040.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220–244.

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

The oblique effect and three-dimensional shape

Elias H. Cohen and Qasim Zaidi

Department of Vision Sciences, State University of New York-State College of Optometry, New York, USA

The classical *oblique effect* refers to the finding that observers are faster and more accurate in discriminating the orientation of a line or edge when it is at or near vertical or horizontal than when it is at an oblique orientation (Appelle, 1972; Mach, 1861). Based on the finding that observer sensitivity to orientation of simple symmetric shapes like an ellipse or the letter "X" also exhibits an oblique effect, Li and Westheimer (1997) suggested that the effect does not arise solely from inequality of simple orientation-tuned receptors in early visual processing, but also involves later orientation processing that can encompass more complex inputs such as shape axes. In this work, we examined how the oblique effect impacts three-dimensional shapes defined by texture cues.

Please address correspondence to Elias H. Cohen, SUNY College of Optometry, Vision Sciences, 33 W. 42nd Street, New York NY, 10036, USA. E-mail: ecohen@sunyopt.edu

Stimuli were upright and obliquely oriented convex and concave wedges (see Figure 1). The texture pattern consisted of sine-wave gratings of the same three spatial frequencies, oriented at 90, $\pm 67.5, \pm 45$, and ± 22.5 degrees with respect to the 3-D axis, and added in randomized phases. The shapes were projected in perspective, and presented at the correct distance and height to be viewed monocularly. In Experiment 1, we varied either the orientation (Experiment 1A) or the location (Experiment 1B) of the axis in small increments around the 90 and 45 degree diameters of the disk. All shapes presented to an observer had the same simulated depth, set at twice the minimum depth required to discriminate convex from concave at any orientation. Observers judged whether the axis was rotated clockwise or counterclockwise in relation to an on-screen reference (two small circles at the stimulus edge indicating a vertical or oblique orientation), or if the axis was shifted to the right or left of centre (i.e., the location of central peak with respect to the reference).

On average, the angular threshold for orientation discrimination around the vertical orientation was 62% lower than around the oblique orientation ($SE = 13\%$). Similarly the average distance threshold for location discrimination of the vertical shape was 53% lower than for the oblique shape ($SE = 9\%$). Thus, the results indicate that sensitivity for 3-D texture-defined shape mirrors that of the traditional oblique effect for 2-D stimuli.

Is this a 3-D effect or is it based on 2-D image properties? Li and Zaidi (2000, 2004) showed that the perception of 3-D shapes from texture cues depends critically on the visibility of orientation flows parallel to the axis of
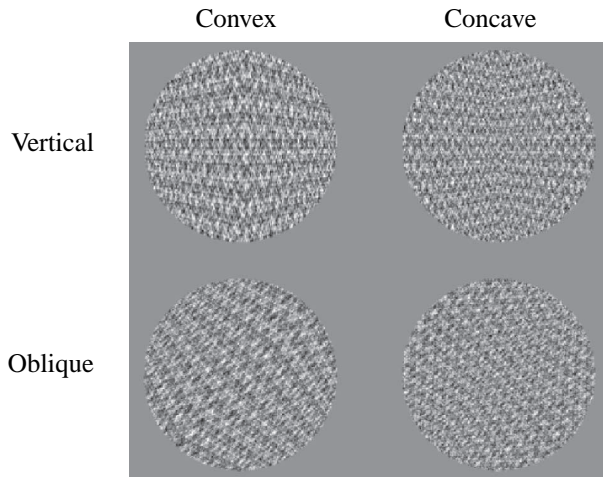

Figure 1.   Demonstrates representative stimuli of both shape types (convex and concave) at the two presented orientations.

maximum curvature, i.e., orthogonal to the 3-D shape axis. These orientation flows are apparent in Figure 1. For example, in the vertical shapes, they converge from left to right for right-slanted 3-D segments, and from right to left for left-slanted 3-D segments. Figure 1 also illustrates that the orientation flows for the vertical 3-D shape vary around the horizontal axis of the image, whereas the orientation flows for the oblique shape vary around an oblique axis of the image. Therefore, before concluding that sensitivity to orientation of 3-D shapes is inherently superior for vertical than oblique orientations, we sought to determine if decreased sensitivity to 2-D oblique orientation flows weakens the percept of an oblique 3-D shape. In that case, the axis discrimination inequality could then arise from the imprecision of the weakened percept.

For this purpose, in Experiment 2, we first measured the minimum 3-D depth required to discriminate convex from concave shapes. Observers judged whether each stimulus was concave or convex, as shapes were varied in depth in a method of constant stimuli. The threshold depth for vertical shapes was 40% lower than the threshold for oblique shapes ($SE = 10\%$). This oblique effect reflects the classical 2-D oblique effect as manifested in detecting small orientation changes from linearity (Westheimer, 2003).

In Experiment 3, the orientation and location discrimination tasks of Experiment 1 were repeated except that, for each subject, 3-D shapes for each orientation were set at twice their discrimination threshold depth. The oblique shapes were thus of much greater simulated depth than the vertical shapes. We reasoned that if 3-D shapes were equated for 2-D oblique effects, differences in performance could be attributed to identification of the 3-D axis. If there is an inherent 3-D oblique effect, we would expect observers to always perform better for the vertical than the threshold-matched oblique shapes. Results from the threshold-matched experiments showed that performance for the vertical condition was no better than the oblique (it was, in fact, worse on average). The oblique effect for 3-D shapes defined by texture cues can thus be attributed entirely to early orientation processing inequalities.

CONCLUSIONS

This study shows that there exists a substantial oblique effect for discrimination of 3-D shapes defined by texture cues, but that it can be completely explained by differences in sensitivity to the different 2-D orientations in texture flows, rather than sensitivity to the orientation of the 3-D axis. An interesting ancillary finding is that observer sensitivity to sign of 3-D curvature is far superior than sensitivity to 3-D orientation or location. Subjects were able to reliably determine shape type for very shallow shapes,

but increased depth was necessary to accurately judge shape orientation or location. A model of 3-D shape extraction built on template-matching to critical orientation flows could explain the discrimination of convexity from concavity despite uncertainty about other shape properties.

REFERENCES

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin*, *78*(4), 266–278.

Li, A., & Zaidi, Q. (2000). The perception of 3D shape from texture is based on patterns of oriented energy. *Vision Research*, *40*, 217–242.

Li, A., & Zaidi, Q. (2004). Three-dimensional shape from non-homogeneous textures: Carved and stretched surfaces. *Journal of Vision*, *4* (10), 860–878.

Li, W., & Westheimer, G. (1997). Human discrimination of the implicit orientation of simple symmetrical patterns. *Vision Research*, *37*, 565–572.

Mach, E. (1861). Über das Sehen von Lagen und Winkeln durch die Bewegung des Auges. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften*, *43*(2), 215–224.

Westheimer, G. (2003) Meridional Anisotropia in visual processing: Implications for the neural site of the oblique effect. *Vision Research*, *43*, 2281–2289.

Testing the relatability hypothesis: Inducer offset, not turning angle, is critical for visual interpolation

Jacqueline M. Fulvio

Department of Psychology, New York University, New York, USA

Manish Singh

Department of Psychology and Center for Cognitive Science, Rutgers University, New Brunswick, NJ, USA

Laurence T. Maloney

Department of Psychology, and Center for Neural Science, New York University, New York, USA

Perceptual organization of cluttered scenes requires identification and grouping of image fragments and their perceptual completion. Despite

Please address correspondence to Jacqueline M. Fulvio, Department of Psychology, New York University, Rm. 877C, 6 Washington Place, New York, NY 10003, USA. E-mail: jmf384@nyu.edu

extensive research, little is known about geometric constraints on visible fragments that permit or preclude completion.

RESEARCH QUESTION

Kellman and Shipley (1991) proposed the notion of "inducer relatability" to predict when visual completion will occur. This notion incorporates two constraints on pairs of contour fragments ("inducers"). The visual system will interpolate a smooth contour between two inducers only if (a) the turning angle between them is $\leq 90°$ and (b) their straight-line *extensions* intersect. In the current study, we manipulated the turning angles between inducers and their relative offsets to test the relatability hypothesis, and to examine the relative importance of its proposed constraints.

GENERAL METHODS

In previous work, we investigated visual *extrapolation* of curved contours by obtaining measurements of the perceived location and gradient of the visually extended contour at multiple distances from the point of occlusion (Singh & Fulvio, 2005). In this study, we extend this method to study *interpolation*. Observers viewed displays (Figure 1A) containing two straight-line inducers that were partially occluded by a rectangle. On each trial, a narrow vertical slit, or "interpolation window", appeared within the occluder at one of six horizontal locations. A straight-line probe was visible through the window. Observers iteratively adjusted the position and orientation of the probe to optimize the percept of a smooth contour connecting the two inducers. We manipulated the turning angle between the inducers (Experiments 1 and 2) and their relative vertical offset (Experiment 2). Some pairs were relatable, some were nonrelatable, and some were "just-relatable" (see details below).

Two experienced observers and one author (JF) made a large number of paired settings of location and orientation in both experiments. To test the influence of turning angle and vertical offset on visual completion, we measured the extent to which observers' settings for a given pair of inducers were consistent with a single smooth contour. This was done by examining the level of internal consistency between their location and orientation settings. We first fit a polynomial to the position data alone, and derived the *predicted* orientation settings based on the polynomial's tangent field. Then we compared the observed orientation settings to these predicted orientation settings. Low values of R^2 indicate that observers' settings do not conform to any single contour—and hence a weak visual interpolation; whereas high values indicate a single perceptually salient interpolated contour.
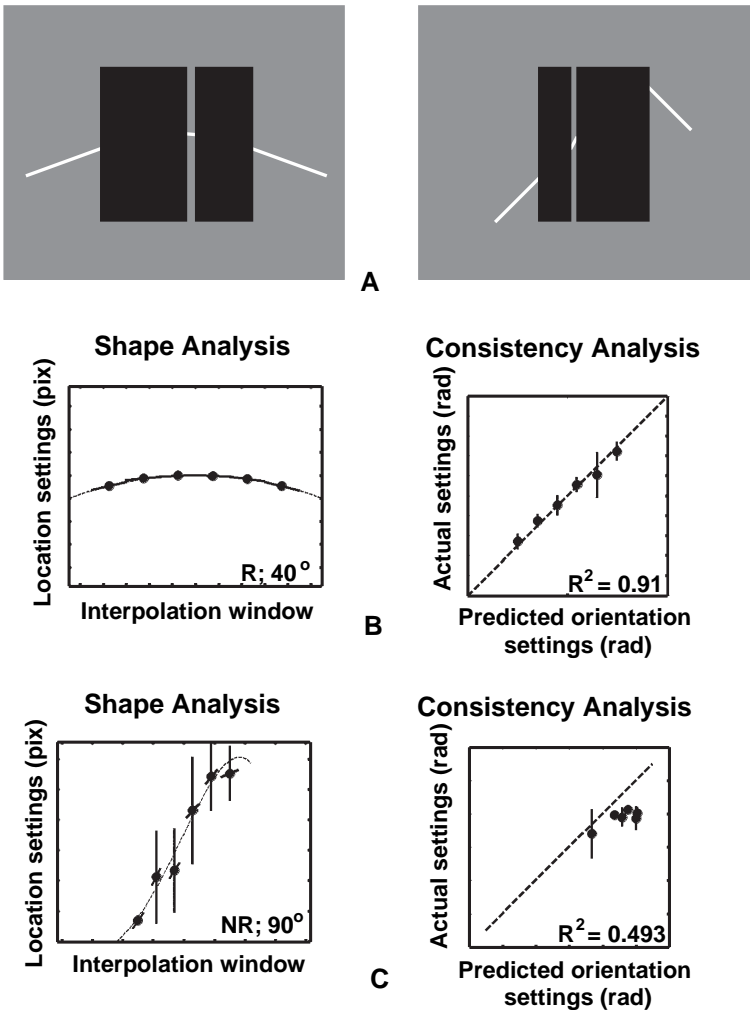
Figure 1. (A) Example stimulus displays from Experiment 1 (left; $=40°$ turning angle condition) and Experiment 2 (right; NR $=90°$ condition). (B) Sample results from Experiment 1 including a parabolic fit to O1's location data in the $=40°$ turning angle condition (left) and the consistency comparison between predicted and observed orientation settings for the same observer in the same condition (right). (C) Sample results from Experiment 2 including a quintic polynomial fit to O3's location data in the NR, $=90°$ turning angle condition (left) and the consistency comparison between predicted and observed orientation settings for the same observer in the same condition (right).

EXPERIMENT 1

We tested the turning angle criterion of relatability using six inducer pairs that varied in turning angle. The turning angles used were 40°, 60°, 80°, 100°,

120°, and 140°. Thus, the first three were relatable and the last three were nonrelatable. The relative heights of the inducers at their points of occlusion were always equal, so that their extensions always intersected, thus satisfying the second criterion of relatability.

The relatability hypothesis predicts that the nonrelatable conditions will not evoke visual interpolation. This should result in a low degree of internal consistency in observers' settings. An example of the best-fit polynomial (parabola) to one observers' location data for the 40° turning angle condition and the corresponding consistency analysis comparing the predicted and obtained orientation settings are shown in Figure 1B. The degree of the polynomial fits to the position data increased with increasing turning angle (the degree of fits ranged from 2 to 6). However, there was no systematic effect of inducer pair *relatability.* Analyses of internal consistency between location and orientation settings revealed high consistency under all conditions, relatable and nonrelatable alike (R^2 values were all ≥ 0.79). Thus, we do not find evidence that choice of turning angle can preclude visual interpolation.

EXPERIMENT 2

We tested the inducer extension intersection criterion of relatability. Since the relative vertical offset (denoted Δ) between two inducers determines whether their extensions will intersect, we manipulated this offset to study the role of the intersection criterion. We began with inducer pairs that had one of two relatable turning angles: 60°, 90°. Applied to each was one of three values of Δ: 0 (as in Experiment 1), $(2w/3)\tan\theta$, and $(4w/3)\tan\theta$, where w is the width of the occluder and θ is the inducer orientation (relative to horizontal). $\Delta = 0$ conditions are relatable, $\Delta = (2w/3)\tan\theta$ conditions are "just-relatable" as the extension of one inducer intersects the endpoint of the opposite inducer, and $\Delta = (4w/3)\tan\theta$ conditions are nonrelatable.

As in Experiment 1, the relatability hypothesis predicts that the nonrelatable conditions will not evoke interpolation, and hence will yield poor setting consistency. The order of the best-fitting polynomial increased with increased offset, Δ (the degree of fits ranged from 2 to at least 5). Here, a systematic effect of inducer pair relatability *was* observed. A high level of internal consistency was obtained in the relatable conditions (R^2 values were all ≥ 0.95), with a decline for the just-relatable (R^2 values were ≥ 0.77) and nonrelatable inducers (R^2 values fell as low as 0.42). An example of the best-fit polynomial (degree $= 5$) to one observer's location data for the NR, 90° turning angle condition and the corresponding consistency analysis (comparing the predicted and obtained orientation settings) are depicted in Figure 1C. The inducer extension intersection criterion of relatability is therefore critical for visual interpolation.

CONCLUSION

We studied the geometric relations between inducers that lead to the percept of a single visually interpolated contour. Specifically, we tested the influence of the two constraints that make up the notion of inducer relatability, by manipulating (a) the turning angle, and (b) the relative offset, between inducers. Using as our measure the internal consistency between observers' position and orientation settings at various locations along the interpolated contour, we found that the intersection constraint (determined by inducer offsets) plays a substantially more important role than the turning-angle constraint. Inducer pairs that violate the intersection constraint (and thus, can be interpolated only with an inflected curve) do not evoke the percept of a single, consistent, interpolating contour.

REFERENCES

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Singh, M., & Fulvio, J. M. (2005). Visual extrapolation of contour geometry. *Proceedings of the National Academy of Sciences USA*, *102*, 939–944.

Attentional orienting in response to peripheral cues survives the attentional blink

S. M. Shahab Ghorashi

University of British Columbia, Vancouver, Canada

Vincent Di Lollo

Simon Fraser University, Burnaby, BC, Canada

Raymond M. Klein

Dalhousie University, Halifax, NS, Canada

If two targets are presented in rapid sequence, the first is usually perceived easily and accurately. Identification of the second target, however, is impaired while the system is busy processing the first target. This second-

Please address correspondence to S. M. Shahab Ghorashi, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC V6T 1Z4, Canada. E-mail: ghorashi@psych.ubc.ca

target deficit, known as the attentional blink (AB; Raymond, Shapiro, & Arnell, 1992), is most pronounced at short intertarget lags and decreases rapidly, vanishing at lags beyond about 700 ms.

A question asked in several earlier studies is whether all aspects of the blinked items are lost, or whether some information is preserved during the period of inattention that is the hallmark of the AB. It has been shown that one's own name (Shapiro, Caldwell, & Sorensen, 1997), word meanings (Maki, Frigen, & Paulson, 1997), and some semantic aspects of words (Rolke, Heil, Streb, & Hennighausen, 2001) can survive the AB. In the present study, we asked whether spatial information about a location cue—presented after the first target but before the second target—can survive the AB.

To this end, we used a rapid serial visual presentation (RSVP) of distractor items consisting of black upper-case letters. Inserted in the stream of distractors were a white letter (first target) and a search array (second target). The search array contained 11 rotated "L"s on an imaginary clock face, which served as distractors in the search task, and the letter "T"—tilted either to the left or to the right—which served as the search target (see Figure 1). The observers' tasks were to report the white letter and then to decide whether the T was tilted to the left or to the right. The search array could appear at Lag 1, Lag 3, or Lag 7 (90, 270, or 630 ms after the onset of the first target). In the cue condition, a small square patch signalled the location of the tilted "T". No cue was presented in the no-cue condition. When present, the cue always appeared at the same time as the item preceding the search display (either the first target or a distractor).
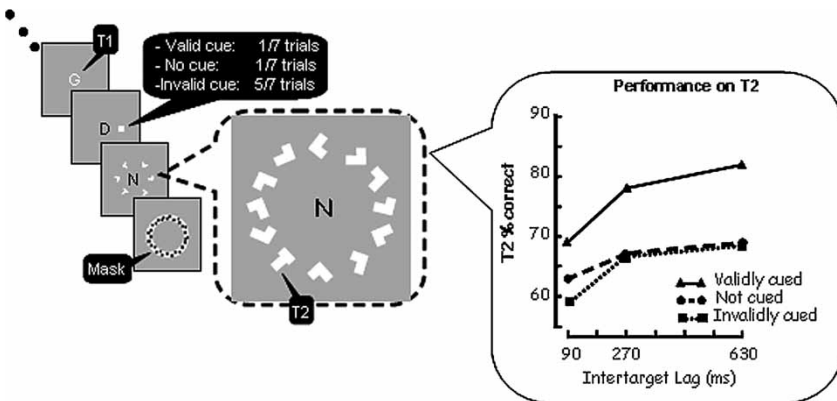

Figure 1. Schematic diagram of the sequence of events in Experiment 3. Experiments 1 and 2 had the same procedure except that the cue was 100% informative and was always present. The graph shows the results for Experiment 3. The Condition × Lag interaction was not significant, confirming that the cue, when valid, survived the AB.

In Experiment 1, an analysis of variance of the percentages of correct responses to the second target revealed a significant effect of lag, indicating an AB deficit, and a significant effect of cue/no-cue (32.3% difference in accuracy, averaged across lags), indicating that the cue facilitated performance on the second target. Importantly, there was no significant interaction between lag and cue/no-cue, indicating that the facilitatory effect of the cue was the same across all lags. This strongly suggests that information about the cue survived the attentional blink. Had perception of the cue been impaired during the period of the AB, the cueing effect should have been muted at the shorter lags, thus producing a significant Lag $\times$ Cue/No-cue interaction effect, which was clearly not the case.

These results, however, gave reason for some concern because in the no-cue condition second-target accuracy was close to chance level, raising the possibility that a floor effect could have prevented the Lag $\times$ Cue/No-cue interaction from occurring. We checked this possibility in Experiment 2, by using critical exposure duration of the second target instead of accuracy as the dependent variable. Specifically, we implemented a dynamic threshold-tracking procedure (PEST; Taylor & Creelman, 1967), in which second-target exposure duration was dynamically adjusted for each participant throughout the experiment, so as to converge on an accuracy level of approximately 80%.

An analysis of the critical durations yielded by the PEST procedure confirmed the results of Experiment 1: there were significant effects of lag (53 ms difference between lag 1 and lag 7 in the cue condition; 123 ms difference in the no-cue condition) and of cue/no-cue (an average of 462 ms difference between the two conditions), but no significant interaction between the two factors. As was the case for Experiment 1, the results of Experiment 2 confirmed that the spatial cue had the same large effect at all lags. In other words, the spatial cue was processed to the same extent during the period of the AB as beyond the period of the AB.

In Experiments 1 and 2, the cue was invariably informative. This raises the possibility that the cue may not, therefore, have been strictly "exogenous". To examine this possibility, we conducted Experiment 3 which was the same as Experiment 1, except that the cue was noninformative. Namely, on any given trial, the target in the search array could be validly cued (in one-seventh of the trials), invalidly cued (in five-sevenths of the trials), or uncued (in one-seventh of the trials). An analysis of the results illustrated in Figure 1 revealed that the accuracy of second-target identification was higher in the validly cued condition than in either the invalidly cued or the uncued conditions, which did not differ significantly from one another. Critically,

the Condition × Lag interaction effect was not significant, confirming that the cue, when valid, survived the AB.

Collectively, these results show that the information needed for exogenous orienting survives the AB (see also Olivers, 2004). The next step in this research programme will be to test whether an "endogenous" cue can escape the AB. Our expectation is that since endogenous cues require identification (as distinct from mere detection, as in the case of exogenous cues), and identification is known to suffer during the AB, we should see an interaction between lag and cue/no-cue conditions. This is because perception of endogenous cues would be impaired at short lags (i.e., during the period of the AB) but not at long lags, with correspondingly less accurate second-target identification at the shorter lags.

It is plausible to suppose that exogenous cues are handled through a process that Di Lollo, Smilek, Kawahara, and Ghorashi (2005) have called "direct matching". Functionally, direct matching is similar to the conventional "preattentive" processing. The critical difference being that in the direct-matching model, the hard-wired modules postulated in the preattentive hypothesis are replaced by programmable input-filtering mechanisms. Future research will examine whether all stimuli that are processed through direct matching are immune from the AB.

REFERENCES

Di Lollo, V., Smilek, D., Kawahara, J., & Ghorashi, S. M. S. (2005). System reconfiguration, not resource limitation, determines the efficiency of visual search. *Perception and Psychophysics*, *67*, 1080–1087.

Maki, W. S., Frigen, K., & Paulson, K. (1997). Associative priming by targets and distractors during rapid serial visual presentation: Does word meaning survive the attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1014–1034.

Olivers, C. L. (2004). Blink and shrink: The effect of the attentional blink on spatial processing. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 613–631.

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

Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology*, *38*, 165–174.

Shapiro, K. L., Caldwell, J., & Sorensen, R. E. (1997). Personal names and the attentional blink: A visual "cocktail party" effect. *Journal of Experimental Psychology: Human Perception and Performance. 23*, 504–514.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, *41*, 782–787.

The landmark effect in perceived object stability: A general mechanism

J. Stephen Higgins, Ranxiao Frances Wang, and David E. Irwin

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

As we interact with our environment we constantly have to update the locations of objects in our visual field. Every time some sort of distraction occurs, like a saccade, a blink, or an occlusion, we have to reevaluate the absolute and relative locations of the objects we are attending to. Prior research has shown that we often make errors if an object is moved during a saccade or a blink. In the case of saccadic suppression of image displacement (SSID), people frequently fail to detect that a target object has moved during the saccade, and in fact the object is erroneously seen as being stable (Deubel, Schneider, & Bridgeman, 1996). It seems as if precise spatial information is not transferred across the saccade and that instead, the visual system makes an assumption of object stability unless the target object is displaced by a great extent.

However, it seems that there must be precise spatial information located somewhere in the brain, because under certain circumstances our motor system is able to make very precise responses across a saccade in response to a small target displacement (Prablanc & Martin, 1992). There does seem to be at least one way of bringing the displacement information to perceptual awareness though. Deubel and colleagues discovered that SSID could be eliminated if presentation of the displaced target object was delayed by 80 or more ms after the end of a saccade. That is, if the target object wasn't present immediately at the end of the saccade people were able to correctly detect the target's movement. They termed this the "blanking effect" (Deubel et al., 1996). We recently demonstrated that suppression of image displacement as well as the elimination of image displacement suppression occurs for eye blinks as well as for saccades (Higgins, Irwin, Wang, & Thomas, 2006). These results suggest that the visual system, as a default, believes that the world is stable and that the only time this assumption is broken is when the target object is not present at the end of the saccade or blink—allowing the system to access the previously stored precise spatial location information. We also determined that this phenomenon was specific to ocular behaviour and not to a general updating mechanism across any visual distraction.

Please address correspondence to J. Stephen Higgins, University of Illinois Urbana-Champaign, 603 E. Daniel Drive, Champaign, IL 61820, USA. E-mail: higgins3@uiuc.edu

Deubel and colleagues found a different pattern of results if a second object, in addition to the target object, appeared in the visual field. They discovered that whenever an object, even if it was not the target object, was presented prior to a second object, the first object was perceived as stable and the second was seen as displaced, regardless of whether it moved or not (Deubel et al., 1996). We sought to determine whether this phenomenon is specific to saccades, or whether it is a more general principle of spatial localization based on temporal object presentation.

METHODS

In this study participants viewed (Figure 1a) a dot and a line for 750 ms, which were $1.76°$ away from each other vertically and $0.47°$ horizontally. The items disappeared for 1200 ms during which there was a blank screen. Either the dot or the line reappeared at this point and then 100 ms later the other item reappeared. Either the first or second item was displaced from its original position by $0.35°$ or $0.71°$. This created four conditions: top item
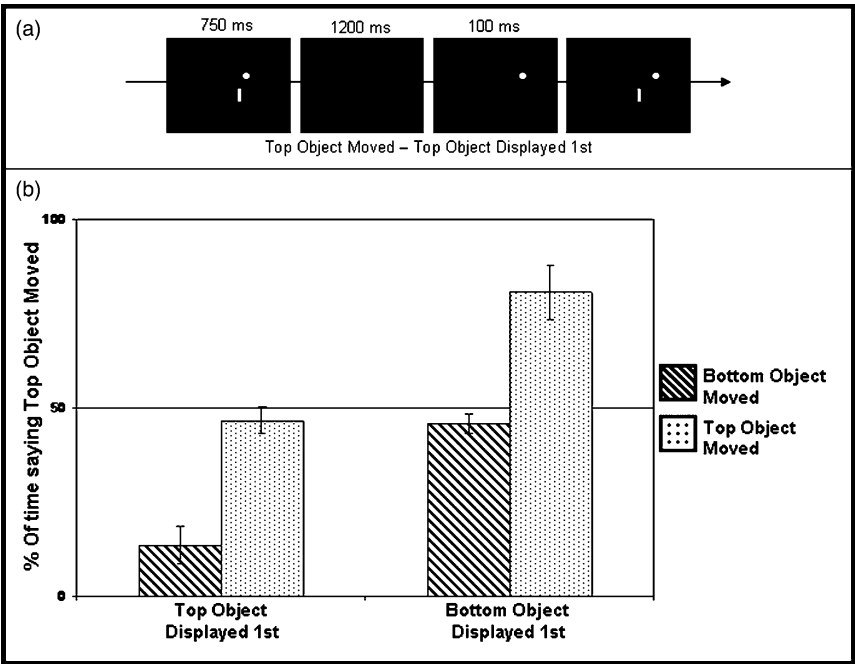

Figure 1. (a) Example of stimuli. (b) When the top object was displayed first it was more often perceived as stable, whether it had moved or not.

appearing first—not moved; top item appearing first—moved; bottom item appearing first—not moved; and bottom item appearing first—moved. Participants reported which item moved from its original position and in which direction it had moved.

RESULTS

Percentage of trials that participants responded that the top item had moved for each of the conditions is presented in Figure 1b. A 2×2 ANOVA showed a significant main effect of both presentation order and actual position moved. These results were also confirmed (not shown) in the opposite response profile (i.e., when the percentage of time people responded that the bottom item had moved). These results indicate that whatever item is shown first is seen as the stable object more often, regardless of whether it actually moved or not. It is also clear that participants had some sensitivity to the actual movement of the items, demonstrated in Figure 1b by the differences between the bars representing the actual target movements. In an eye tracking version of this experiment, which eliminated all trials in which an eye movement was made between the presentation of the first set of items and their re-presentation, we confirmed that eye movements had no effect on the pattern of results.

CONCLUSIONS

It has been suggested that at the end of a saccade people seek out a reference point to determine the relationship of the objects around it, and to decide whether something has moved from its original position or not (e.g., Deubel et al., 1998). Our results indicate that this is true not only after saccades, but can be generalized to object updating after any generic distraction. The results of this study indicate that when re-presented with more than one item, separated by a temporal interval, people tend to use object relative comparisons to determine which object has moved, rather than relying on absolute spatial positions.

REFERENCES

Deubel, H., Bridgeman, B., & Schneider, W. X. (1998). Immediate post-saccadic information mediates space constancy. *Vision Research*, *38*, 3147–3159.
Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research*, *36*(7), 985–996.

Higgins, J. S., Irwin, D. E., Wang, R. F., & Thomas, L. E. (2006). *A blink is not a blank, but it can be a saccade*. Paper presented at the annual meeting of the Psychonomic Society, Houston, TX.

Prablanc, C., & Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *Journal of Neurophysiology*, *67*, 455–469.

The role of familiarity in 3-D view transferability of identity adaptation

Fang Jiang

School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, TX, USA

Volker Blanz

Fachgruppe Medieninformatik, University of Siegen, Siegen, Germany

Alice J. O'Toole

School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, TX, USA

High-level face adaptation effects show substantial tolerance to two-dimensional affine transformations in size, orientation, and retinal position (e.g., Leopold, O'Toole, Vetter, & Blanz, 2001; Watson & Clifford, 2003; Zhao & Chubb, 2001). More recently, partial transfer of face aftereffects over three-dimensional view changes has also been reported (Jeffery, Rhodes, & Busey, 2006; Jiang, Blanz, & O'Toole, 2006). For example, Jiang et al. (2006) found that identity adaptation transferred across a 30° change in viewpoint, although the size of aftereffects was diminished in comparison to the aftereffects found when the adapting and test faces were presented from the same viewpoint. Partial three-dimensional transfer was also reported by Jeffery et al. (2006) using figural shape aftereffects (e.g., Webster & MacLin, 1999). They found significant, but reduced aftereffects, when the adapting and test views differed by 45 degrees.

One factor not considered in previous studies is the role of familiarity in adaptation transfer. Psychophysical and neurophysiological studies indicate

that representations of objects and faces become more robust and generalizable with increased familiarity (e.g., Booth & Rolls, 1998; Burton, Bruce, & Hancock, 1999). We manipulated participants' experience with faces and measured identity adaptation effects within and across viewpoint. We hypothesized better transfer of adaptation for more familiar faces.

Face stimuli were generated using a three-dimensional morphable model (Blanz & Vetter, 1999), implementing a multidimensional prototype-centred face space (cf. Jiang et al., 2006). Four male faces served as original faces. Antifaces were created with -0.75 identity strengths, avoiding morphing artefacts at more extreme identity strengths. Anticaricatures were generated at 0.10 and 0.35 identity strength levels. Rotated faces were created by rotating faces 30 degrees to the right.

FAMILIARIZATION

Participants were familiarized with four original faces and their 0.35 level anticaricatures. The exposures and views of the familiarization stimuli varied in six familiarity conditions. In the low familiarity (LF) condition, participants saw each frontal-view familiarization face twice. In the medium familiarity (MF) condition, participants saw each frontal-view familiarization face four times. In the high familiarity (HF) condition, the total number of exposures was set to eight. The HF condition was further divided into three familiarization conditions in which participants saw only frontal views (HFFV), only rotated views (HFRV), or half frontal half rotated views (HFMV), respectively. Finally, in the extreme familiarity (EF) condition, participants saw each face 16 times from frontal view and 16 times from rotated view.

PRACTICE

Participants were tested on 40 practice trials in the format of test trials, but with 0.35 anticaricatures as test faces.

TEST

For all three adaptation conditions, participants identified briefly (200 ms) flashed, frontal views of 0.10 level anticaricatures as one of the four original faces. In no adaptation trials, participants identified anticaricatures without adaptation. In within-view adaptation trials, the identification task was performed following 5 s of adaptation to a frontal antiface. In across-view adaptation trials, the identification task was performed following 5 s of adaptation to a rotated antiface.

A total of 102 volunteers participated in the study, with a minimum of 15 in each of the six familiarity conditions. Identity adaptation effects occurred

both within and across viewpoint, $F(2, 168) = 262.59$, $p < .0001$ (see Figure 1), replicating previous findings (Jiang et al., 2006). The magnitude of adaptation effects in both within- and across-view conditions increased with familiarity, $F(5, 84) = 6.68$, $p < .0001$ (see Figure 1). There was a marginal interaction between adaptation condition and familiarity level, $F(10, 168) = 1.78$, $p = .068$, although a direct interpretation of this is not possible due to the qualitative and quantitative differences among the familiarity conditions. This marginal interaction could be driven by the fact that the magnitude of across-view adaptation effects approached the magnitude of within-view adaptation effects in the EF condition.

Planned comparisons were used to test for differences among three high familiarity conditions. There was no advantage for familiarity with rotated view over familiarity with frontal view (HFRV > HFFV), $F < 1$, nor was there a significant advantage for familiarity with multiple views versus familiarity with frontal or rotated view alone (HFMV against HFFV and HFRV), $F(1, 84) = 2.28$, $p = .17$.

The findings of this study are as follows. First, familiarity strengthens the overall magnitude of the identity aftereffect. Experience with a face, therefore, contributes to the development of a more robust and malleable representation. Second, familiarity enhances the view transferability of identity adaptation. This finding has implications for understanding how we build a view-independent neural representation of faces. The data suggest that a more robust representation, which is less vulnerable to changes in viewing conditions, can evolve through experience.
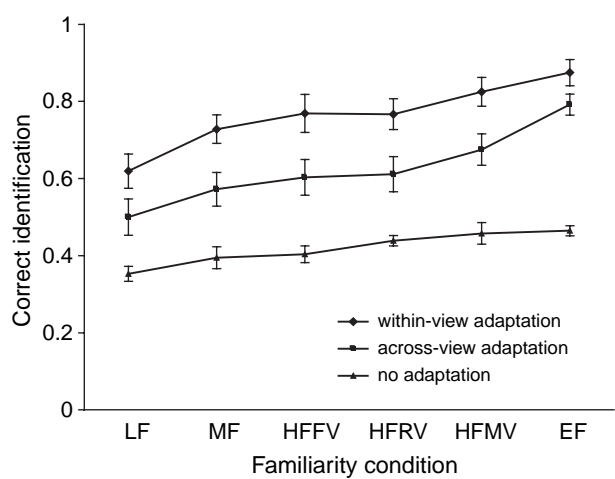

Figure 1. Effects of familiarity on identity adaptation. The within- and across-view adaptation lines show the proportion of correct identification in these respective conditions as a function of the familiarity condition. The no-adaptation line indicates identification accuracy without adaptation.

Although we failed to find a significant advantage for familiarity with multiple views over familiarity with a single view, only two views were included in our multiple view condition. It is possible that the information gained from a single rotated view was not sufficient to yield a multiple view condition advantage. This issue can be addressed in future studies that systematically manipulate the familiarization views.

In summary, identity adaptation provides a useful tool for probing real face representations as they evolve.

REFERENCES

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of three-dimensional faces. In *SIGGRAPH '99 Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 187–194). New York, NY: ACM Press/Addison-Wesley Publishing Co.

Booth, M. C. A., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex*, *8*, 510–523.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From pixels to people: a model of familiar face recognition. *Cognitive Science*, *23*, 1–31.

Jeffery, L., Rhodes, G., & Busey, T. (2006). View-specific coding of face shape. *Psychological Science*, *17*, 501–505.

Jiang, F., Blanz, V., & O'Toole, A. J. (2006). Probing the visual representation of faces with adaptation: A view from the other side of the mean. *Psychological Science*, *17*, 493–500.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*, 89–94.

Watson, T. L. & Clifford, C. W. G. (2003). Pulling faces: An investigation of face-distortion aftereffect. *Perception*, *32*, 1109–1116.

Webster, M. A., & MacLin, O. H. (1999). Figural after-effects in the perception of faces. *Psychonomic Bulletin and Review*, *6*, 647–653.

Zhao, L., & Chubb, C. F. (2001). The size-tuning of the face-distortion aftereffect. *Vision Research*, *41*, 2979–2994.

Can capture of visuospatial attention be contingent on category? Behavioural and electrophysiological evidence

Émilie Leblanc and Pierre Jolicoeur

Département de Psychologie, Université de Montréal, Montréal, QC, Canada

Contingent attentional capture is the involuntary allocation of attention to an item that shares the target-defining property, even though it is irrelevant

Please address correspondence to Émilie Leblanc, Département de Psychologie, Université de Montréal, C.P. 6128 Succursale Centre-ville, Montréal, Québec, Canada H3C 3J7. E-mail: emilie.leblanc-langlois@umontreal.ca

to the task at hand, because it matches the observer's top-down control settings (e.g., Folk, Remington, & Jonhston, 1992). Such effects have been observed for low-level characteristics such as sudden onset, movement, colour, and shape (e.g., Bacon & Egeth, 1994; Folk & Remington, 1998; Folk et al., 1992). In a recent electrophysiological study, the presence of an N2pc in response to a target-coloured distractor demonstrated that in contingent capture by colour, the capturing distractor generates a shift of visuospatial attention to its location (Leblanc, Prime, & Jolicoeur, 2006). The N2pc component, a greater negativity observed in the N2 time range at posterior electrode sites contralateral to the attended stimulus, is often used as an index of the orienting of visuospatial attention (Luck & Hillyard, 1994).

The present study had two goals. First, we wanted to explore the possibility that attentional control settings could be established in favour of a higher level target attribute, category. We also wished to discover if a capture effect involving category would result from a visuospatial shift of attention to the distractor location, or from interference at later processing stages. We designed a paradigm in which category was the target defining characteristic, and we recorded event-related potentials (ERPs) while subjects performed the task. Once again, we focused on the N2pc component of ERPs to track the locus of visuospatial attention following the presentation of target- and nontarget-category peripheral distractors.

METHODS

The stimuli were grey characters displayed in rapid serial visual presentation at fixation on a black screen, at a rate of 117 ms per item, without any blank intervals between items. For half the subjects, the stream was comprised mainly of letters, and the target, to be reported at the end of the trial, was the only digit present in the stream; for the other half, the distractor and target categories were reversed. Two peripheral distractors flanked the item presented two frames (234 ms) before the target. One distractor was always a "#" sign, and the other was a character. In 50% of the trials, the distractor character shared the target category (target-category distractor condition), and in the remaining 50% of the trials it did not (nontarget-category distractor condition). For example, if targets were digits (e.g., "2"), target-category distractors were other digits, different from the target digit (e.g., "5"), and nontarget-category distractors were letters (e.g., "F").

RESULTS

Mean percentages of correct target identifications for 16 subjects were entered in a repeated measures analysis of variance (ANOVA) with target

category (digit or letter) as a between-subjects factor and distractor category condition (target or nontarget category) as a within-subjects factor. Only the main effect of distractor category condition was significant, $F(1, 14) = 53.04$, $p < .001$, as accuracy was nearly 8% lower when the target was preceded by a target-category peripheral distractor (86.6%) than by a nontarget-category peripheral distractor (94.1%). This result indicates that the presence of a target-category distractor produced contingent attentional capture.

The ERPs were time locked to the distractor onset, and contralateral minus ipsilateral subtraction waveforms were computed for each distractor category condition. A pooling of six electrode pairs (P3/P4, P5/P6, P7/P8, PO3/PO4, PO7/PO8, and O1/O2) was used for the analyses. The obtained subtraction curves seemed to display two distinct negative components contralateral to the distractor location, one occurring in the N1 time range, between 130 and 200 ms postdistractor onset, and the other, in the N2 time range, between 200 and 290 ms postdistractor onset (see Figure 1). Mean amplitudes in these time windows were entered into ANOVAs with distractor category condition as a within-subjects factor. While target- and nontarget-category distractors generated statistically equivalent contralateral negativities in the 130–200 ms time window, $F(1, 15) = 0.86$, $p = .37$, target-category distractors elicited a greater N2pc (200–290 ms time window) than nontarget-category distractors, $F(1, 15) = 4.79$, $p = .04$.

DISCUSSION

The first aim of this study was to assess contingent capture by target category. The behavioural results revealed that such capture did occur, as subjects performed fewer correct target identifications when the target
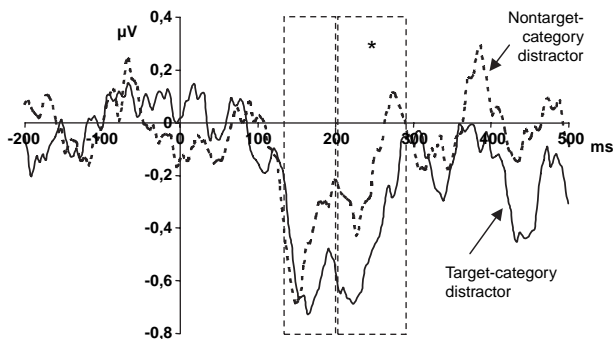

Figure 1.   ERP subtraction waveforms for each distractor category condition.

was preceded by target-category distractors than by nontarget-category distractors.

Our second goal was to determine if the interference produced by distractors that matched the top-down attentional control settings with regards to category reflected a visuospatial effect, as has been found for contingent capture by colour. The present ERP results revealed that the presentation of target-category distractors was associated with a greater N2pc than that of nontarget-category distractors, suggesting that visuospatial attention was deployed to the location of target-category distractors. We also found an earlier contralateral negative component, present in both distractor conditions. This set of findings is very similar to what was observed in a recent study by Hopf, Boelmans, Schoenfeld, Luck, and Heinze (2004). In a visual search task, they found that distractors that matched a target attribute (orientation) that was not the target-defining characteristic (colour) elicited a "relevant-orientation distractors related negativity" over posterior electrode sites contralateral to the distractors, starting approximately 140 ms and lasting until 300 ms poststimuli onset, hence overlapping with the N2pc to the target. The authors interpreted this finding as an early correlate of attentional selection based on target features that preceded the deployment of visuospatial attention to the target location. A similar process might have come into play in the present study. In every trial, one "#" distractor and one character distractor were presented. Because subjects were set to search for a character, letter, or digit, the character matched the target features to a certain point, and therefore could have been selected for further analysis. However, only a distractor that matched the target category attracted visuospatial attention to its location, hence the difference in the N2pc time range between the two conditions. The present results therefore suggest that more than one process may be involved in contingent capture by category, but that it includes a visuospatial component.

REFERENCES

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception and Psychophysics*, *55*, 485–496.

Folk, C. L., & Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 847–858.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030–1044.

Hopf, J.-M., Boelmans, K., Schoenfeld, M. A., Luck, S. J., & Heinze, H.-J. (2004). Attention to features precedes attention to locations in visual search: Evidence from electromagnetic brain responses in humans. *Journal of Neuroscience*, *24*, 1822–1832.

Leblanc, É., Prime, D. J., & Jolicoeur, P. (2006). *Contingent attentional capture has a visuospatial locus: Evidence from human electrophysiology*. Manuscript submitted for publication.

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

The unique contributions of retinal size and perceived size on change detection

Brian R. Levinthal and Alejandro Lleras

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

Is change detection (CD) wholly dependent on low-level features of objects (e.g., magnitude, colour, local contrast), or is performance also modulated by some interpretation of these features in the context of a scene? Ro, Russell, and Lavie (2001) demonstrated that participants were able to detect changes to faces faster than changes in other objects. While this claim has been disputed (Palermo & Rhodes, 2003, argued that it is the result of an "odd one out" bias), it suggests that stimuli are interpreted beyond a simple "pixel-by-pixel" comparison during CD.

How might CD performance be influenced by a scene's context when there is no "odd one out"? Our study investigated whether the addition of pictorial depth cues influenced detection performance by altering the perceived size of objects. There is evidence that placing a target further in a depth plane reduces the time to find the target (Aks & Enns, 1996); however, this issue has not been investigated within the context of CD, nor has previous research directly compared the influences of 2-D and 3-D size. The present study attempts to go beyond those initial findings to directly assess the contribution of "interpreted representations" (here, post size-constancy) on CD performance, above and beyond that of simple low-level features.

Please address correspondence to Brian R. Levinthal, University of Illinois Urbana-Champaign, Apartment 103, 311 W. Clark Street, Champaign, IL 61820, USA. E-mail: levintha@uiuc.edu
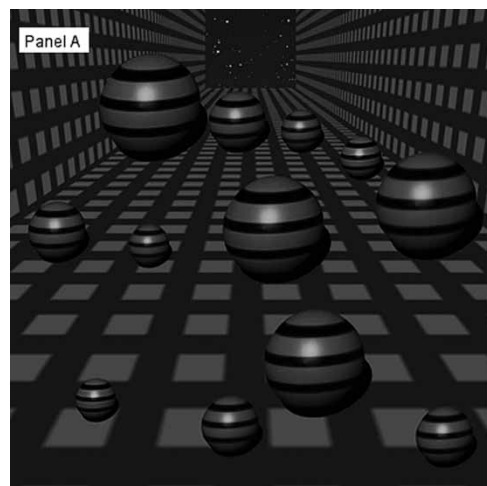
METHODS

Participants viewed a flashing scene consisting of slanted depth-inducing planes and twelve coloured spheres (Figure 1A). Between flashes, the colour of one sphere changed, and the participants' task was to quickly determine which object had changed. Participants wore an eye-patch over their nondominant eye to minimize conflicting binocular depth cues and enhance the perception of depth in the displays.

The images were generated using 3-D modelling software (Blender). The depth-inducing plane consisted of a floor and two wall surfaces (a lattice pattern was applied to the surfaces as an additional depth cue). Spheres were arranged over three separate regions on the floor surface (near, mid, and far). There were three 2-D sizes, and the placement of a sphere on one of the three floor regions resulted in five different 3-D sizes. 2-D sizes and floor regions were selected such that a large sphere in the near-region shared the same 3-D size as a medium-sized sphere in the mid-region and a small sphere in the far-region. Three spheres appeared in each region (one of each 2-D size), and an additional sphere of each 2-D size was positioned randomly between regions (they were never CD targets). Five colours were used, and no colour was ever present on fewer than two spheres or more than three spheres.

The display was presented for 250 ms, after which the spheres disappeared for 250 ms (the depth-inducing background remained on-screen). One sphere changed colour between presentations, which cycled until the participant indicated detection (via button-press). A static greyscale version of the scene then appeared, and participants selected the object that had changed with the mouse. Performance was measured by detection time (from the first presentation until the participants' response) and response accuracy (whether the correct object was selected).

RESULTS

For all participants, response accuracy was high (approximately 95%), therefore our analysis focused primarily on CD time (Figure 1B). As 2-D size increased from smallest to largest, detection time decreased by nearly 1000 ms. Over and above the effect of 2-D size, CD time decreased by approximately 400 ms as 3-D size increased. These results were confirmed by a hierarchical linear model in which the effect of 3-D size was evaluated after controlling for 2-D size. Both effects were significant (each at $p < .001$).

Figure 1. (A) A typical display for this experiment (during the task, the spheres appeared in various colours). (B) As 2-D size increased, detection times decreased; a similar effect was observed as 3-D size increased.
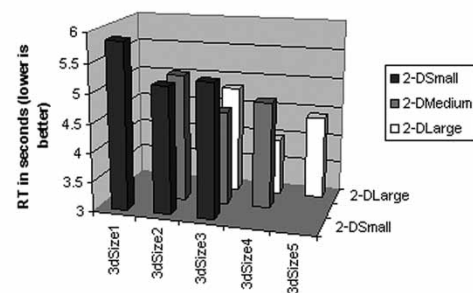
	3dSize1	3dSize2	3dSize3	3dSize4	3dSize5
2-DSmall	5.85	5.16	5.28		
2-DMedium		5.17	4.6	4.81	
2-DLarge			4.82	3.97	4.41

DISCUSSION

Our results show that, after accounting for the influence of 2-D size, there is a significant benefit to CD performance as perceived size increases. While the 2-D size of the object does indeed influence CD, postconstancy representations also seem to influence performance.

To quantitatively compare the effects of 2-D size and 3-D size, we conducted a nulling experiment (Experiment 2) to assess the magnitude of the induced depth illusion. Fifteen participants viewed two spheres against a depth-inducing plane (one positioned in the near-region and one in the far-region) and adjusted the diameter of the more distant sphere until both objects appeared to be the same perceived size. Participants perceived a robust size illusion, overrepresenting the area of the more distant object by 32.4%.

Experiment 2 also examined whether the effect of 3-D size in Experiment 1 reflected a true contribution of 3-D size, rather than an artefact of the stimuli or experimental procedure. Participants completed seven blocks of CD trials (which occurred between two blocks of the nulling experiment) which were identical to the trials of Experiment 1, except that depth-inducing planes were removed. In contrast to Experiment 1, the HLM analysis revealed that 2-D size and "previous 3-D size" (i.e., the 3-D size an object would have been had the depth-inducing planes been presented) no longer independently contributed to CD performance, ruling out the possibility of stimulus artefacts in Experiment 1.

CONCLUSION

While it may be unsurprising that the retinal size of objects influences change detection performance, it is perhaps less obvious that such a finding would extend to the perceived size of objects. Our results suggest that the representations used in a CD tasks are not merely "pixelized" representations of the world. Rather, these representations seem to be "contextualized", such that 2-D images are integrated with depth cues, with the ensuing post size-constancy representations impacting our detection of changes in the scene.

REFERENCES

Aks, D., & Enns, J. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1467–1481.

Palermo, R., & Rhodes, G. (2003). CD in the flicker paradigm: Do faces have an advantage? *Visual Cognition*, *10*(6), 683–713.

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological Science*, *12*, 94–99.

Eye movements and individual differences

Xingshan Li, Keith Rayner, Carrick C. Williams, Kyle R. Cave, and Arnold D. Well

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

Eye movements have been widely used to study cognitive processes (Rayner, 1998). Most of these studies dealt with eye movements in a single task. We are interested in whether the same participant shows the same or different eye movement patterns in different tasks. This kind of study is important to understand the common neural circuitry underlying eye movement control in different tasks. It is also helpful to understand the variability in eye movement data.

Prior research by Andrews and Coppola (1999) found that eye movement patterns were different in active viewing tasks compared to passive viewing tasks. Other studies compared eye movement patterns in reading and visual search (Vitu, O'Regan, Inhoff, & Topolski, 1995; Rayner & Fischer, 1996; see Rayner, 1998, for a review). The current study differs from previous studies in that we employed a wide range of tasks and we also used participants with varying amounts of knowledge of English and Chinese.

In this study, the eye movements of Chinese readers and English readers were monitored during a series of different cognitive tasks (reading, scene perception, visual search, and face perception). One purpose of the study was to examine individual differences in eye movements. Another purpose of this study was to explore how eye movements are influenced by tasks and experiences. Because of differences in terms of the amount of experience in Chinese reading and English reading, we would expect differences in reading tasks among different groups, and perhaps also differences in nonreading tasks involving Chinese characters.

PARTICIPANTS

Four groups of participants were recruited based on their Chinese reading level. Twenty-nine participants attended school in China, learning to read Chinese in elementary school and starting to study English in middle school, so that they are now fluent in both. The three remaining groups all attended school in the USA, and started learning to read English in elementary school. Some had Chinese parents, and others had American parents. Six participants learned to read Chinese in middle or high school; thirteen participants learned to read Chinese at college; and twenty-six participants never studied Chinese. These participants are different in terms of their ability to read English and Chinese, and also in terms of cultural background.

TASKS AND PROCEDURES

All of the participants performed the following six tasks: English reading, face processing, scene perception, visual search, Chinese character count, and Chinese character search. In the English reading task, participants read 40 English sentences. In the face processing and scene perception tasks, the participants were shown 16 pictures of female faces and 24 pictures of scenes, and were asked to remember them for a later memory test. In the visual search task, participants were asked to find a brown square among brown circles and pink squares. The target was present in half of the trials. In the Chinese character count task, the participants counted the occurrences of a Chinese character in a paragraph of Chinese text. In the Chinese character search task, participants searched for a specific character in a randomly arranged array of Chinese characters. The target was also present in half of the trials. Finally, for those participants who could read Chinese, we also asked them to read 36 Chinese sentences. All of these tasks were preceded by detailed instruction in English and some practice trials. Eye movements were monitored using an Eye Link 2 eyetracker.

MEASUREMENTS

How viewers decide when and where to move their eyes is an important question in eye movement research (Rayner, 1998). Previous research has indicated that these two aspects of eye movements are controlled by different brain mechanisms (Aslin & Shea, 1987; Becker & Jürgens, 1979; see Rayner, 1998, for review). We used fixation duration, saccade size, and number of fixations to measure these two aspects of eye movements. It is generally assumed that (1) fixation duration reflects the time needed to process the information in the perceptual span and the time needed to plan next saccade

and (2) saccade size is related to the size of the perceptual span and how the next saccade target is selected.

RESULTS

The eye movement records revealed the following findings. First, we found individual differences in fixation duration in that participants who made longer fixations in one task tended to make longer fixations in other tasks. This pattern was not apparent in the number of fixations and saccade size. Second, in Chinese character count and Chinese character search tasks, Chinese participants made longer saccades and spent less time on each fixation than native English readers. Their experience with these characters made them more skilled at these tasks. Third, and most surprisingly, fixation duration differences between participant groups also appeared during the scene perception task (see Figure 1). Thus, even when no characters from any language were being viewed, Chinese participants spent less time on each fixation. Fourth, in the reading tasks, the fixation duration decreased and saccade size increased with the amount of experience with that language. Finally, we found that across all participant groups, task influenced both fixation duration and saccade length.
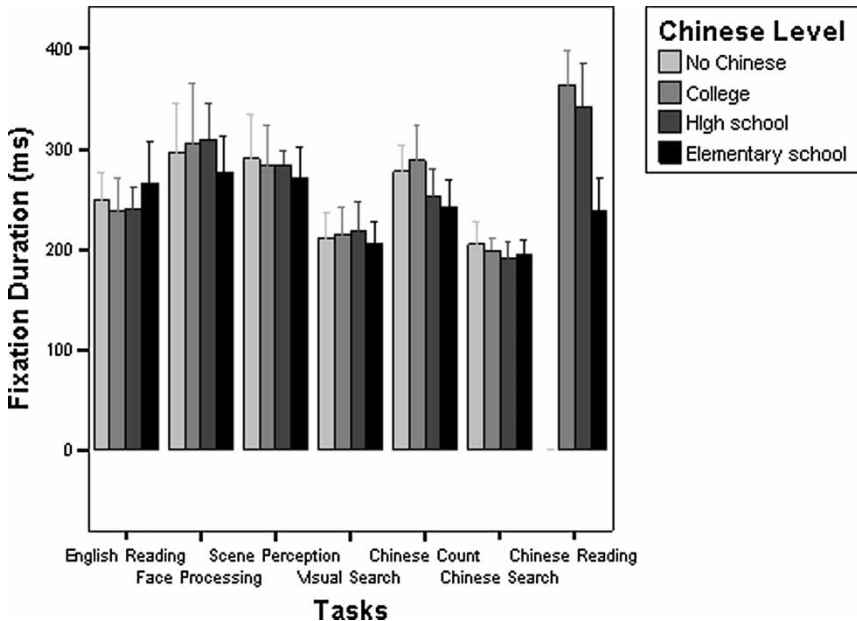

Figure 1.   Fixation duration for the different participant groups for the different tasks.

DISCUSSION

The fact that participants who made longer fixations in one task tended to make made longer fixations in the other tasks suggests that there is a common circuit controlling when to move the eyes in the different tasks. However, the saccade length data did not show this pattern, perhaps because the "where" system depends more on task and stimuli. Consistent with some other studies in reading, increased ability in a language reduced fixation duration and increased saccade size in reading (see Rayner, 1998, for a review). We also found the same pattern in Chinese count and Chinese search tasks. Overall, the eye movement pattern is influenced by all of the factors we examined: individual differences, the nature of the task, and experience with the task.

REFERENCES

Andrews, T. J., & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Visual Research*, *39*, 2947–2953.

Aslin, R. N., & Shea, S. L. (1987). The amplitude and angle of saccades to double-step target displacements. *Visual Research*, *27*, 1925–1942.

Becker, W., & Jürgens, R. (1979). Analysis of the saccadic system by means of double step stimuli. *Visual Research*, *19*, 967–983

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics*, *58*, 734–747.

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eye movement characteristics are similar in scanning letter strings and reading text. *Perception and Psychophysics*, *57*, 352–364.

The control of fixation duration during scene perception

Graham L. Pierce

Michigan State University, East Lansing, MI, USA

John M. Henderson

University of Edinburgh, Edinburgh, UK

Recent research on gaze control during scene perception has primarily focused on the placement of fixations. However, fixations also differ in their

Please address correspondence to Graham L. Pierce, Department of Psychology, Michigan State University, East Lansing, MI 48824, USA. E-mail: glpierce@msu.edu

durations, and duration may be at least as important an indicator of attentional focus as location (Henderson, 2003). In this study we investigated whether the durations of individual fixations during scene perception reflect current visual processing. We tracked participants' eye movements as they viewed photographs of real-world scenes, and we used a saccade-contingent stimulus onset delay (SOD) paradigm to delay the appearance of the currently presented scene at the beginning of specific fixations. The duration of the delay was varied and the influence of the delay on the duration of the current fixation was measured. Direct control models of fixation duration posit that fixations are influenced by currently available stimulus information and predict a monotonic increase in fixation duration as a function of stimulus onset delay. In contrast, indirect control models propose that fixation durations are determined by general oculomotor control settings and therefore predict no influence of immediate delay on current fixation duration.

METHODS

Thirty-six participants each viewed 40 full-colour photographs of natural scenes in preparation for a later memory test. Participants were told that the scenes would occasionally flicker and that they should ignore this as best they could. A colour noise mask delayed scene onset on every tenth fixation. In Experiment 1, SODs of 0, 100, 200, 600, and 1200 ms were used, and trials were presented in two blocks such that 20 consecutive scenes had randomly selected SODs per trial and the remaining 20 scenes used only one SOD per trial (order counterbalanced across subjects). Experiment 2a used randomly selected SODs in each trial of 0 (control), 40, 75, 100, 250, and 450 ms, and Experiment 2b used randomly selected SODs of 150, 200, 300, 400, 600, and 750 ms.

RESULTS AND DISCUSSION

A $5 \times 2 \times 2$ (SOD $\times$ Blocking condition $\times$ Order) analysis of variance was conducted on the data from Experiment 1 to determine whether SOD affected fixation duration differently when it was blocked versus random. While the effect of SOD was significant, $F(4, 100) = 727.28$, $p < .01$, neither blocking condition nor condition order were significant, nor were any interactions involving these factors. The lack of an interaction between SOD and blocking condition suggests that SOD effects were not strategically generated.

Because there was no effect of blocking, the data from Experiments 1 and 2 were combined in the main analysis. The distributions of fixation durations

for SODs greater than 150 ms were bimodal at each SOD (see Figure 1). The two modes at each SOD were entered into separate linear regressions. The slope of the best fitting regression line for the upper set of modes was 1.01 (Fixation Duration $= 252.09 + 1.01 *$ SOD; $R^2 = 1.00$, $p < .01$). These data provide a strong demonstration that during scene perception, one population of fixations is under the immediate and direct control of the current visual stimulus. In contrast, for the lower set of modes, there was no significant slope. These latter data suggest that a second population of fixations is controlled by factors that are independent of the immediate stimulus input.
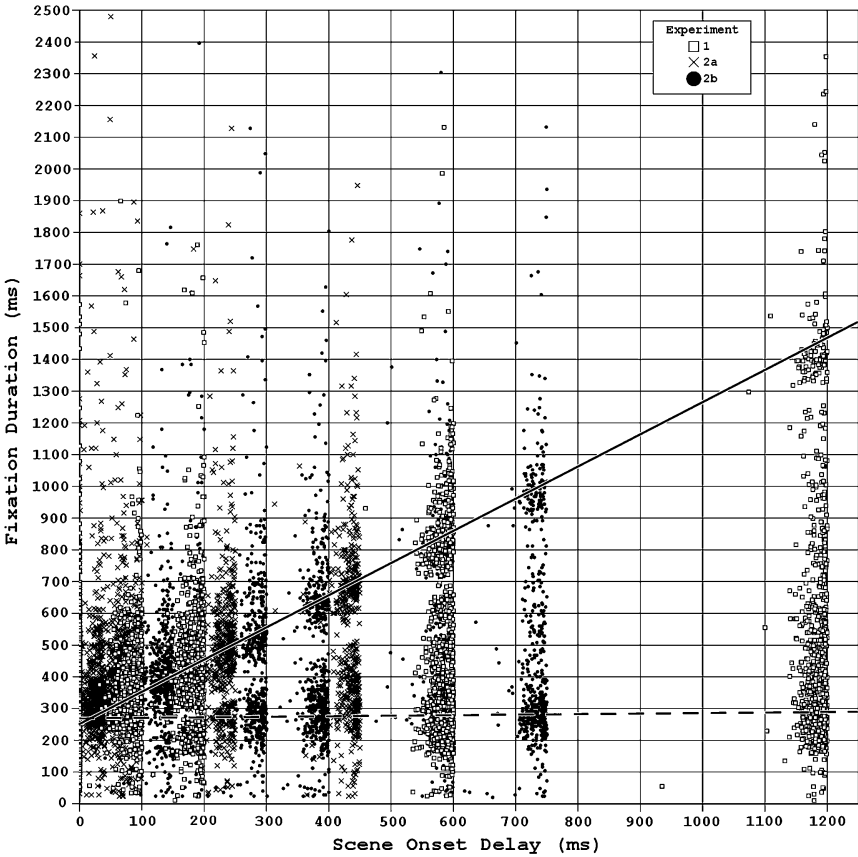

Figure 1.   Fixation duration as a function of scene onset delay in all three experiments. Experiment 1 data are denoted by squares, Experiment 2a by crosses, and Experiment 2b by circles. Fitted linear regression lines are shown for upper (solid) and lower (dashed) modes for SOD bins greater than 150 ms.

Further evidence for direct control was provided by a dramatic decrease in the probability of a fixation terminating at about 100 ms following stimulus onset (seen as the diagonal gap cutting across the fixation distributions in Figure 1). It has been suggested that this saccadic inhibition effect is due to the interruption of saccade programming when a new stimulus appears (Reingold & Stampe, 2002).

CONCLUSIONS

The present study was designed to determine whether the durations of individual fixations during scene perception are under direct stimulus control, are controlled by mechanisms independent of current visual input, or both. We found strong evidence for both direct and indirect control. One population of fixations increased in duration in a one-to-one manner as a function of stimulus onset delay. These data provide a striking demonstration of the influence of the current stimulus on fixation duration. When the availability of useful information is delayed, the gaze control system holds fixation until such information has become available. In contrast, a second population of fixations remained constant over stimulus onset delay. These fixations are apparently controlled by non-stimulus factors such as general oculomotor processes and task-specific global gaze control parameters. The observation of two populations of fixations is in line with the results of SOD studies in reading (Morrison, 1984; Rayner & Pollatsek, 1981) and has several important implications. First, although a great deal of recent computational modelling of gaze control has focused on predicting fixation locations (Itti & Koch, 2001; Torralba, Oliva, Castelhano, & Henderson, 2006), a complete model of gaze control in scene viewing will also have to account for fixation durations. Second, the moment-to-moment dynamics of attention allocation in scenes is reflected by both fixation location and fixation duration. Studies that focus on fixation location alone to generate an "attentional landscape" may produce a mistaken understanding of how attention is allocated over complex scenes. Third, given the close correspondence observed here between a large population of fixation durations and on-line scene processing, fixation durations would seem to offer a powerful noninvasive real-time measure of active scene perception.

REFERENCES

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention, *Nature Reviews Neuroscience*, *2*, 194–203.

Morrison, R. E. (1984). Manipulation of scene onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception & Performance*, *10*(5), 667–682.

Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology*, *33A*, 351–373.

Reingold, E. M., & Stampe, D. M. (2002). Saccadic inhibition in voluntary and reflexive saccades. *Journal of Cognitive Neuroscience*, *14*, 371–388.

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

The world is too much: Effects of array size on the link between language comprehension and eye movements

Dane W. Sorensen and Karl G. D. Bailey

Andrews University, Berrien Springs, MI, USA

The visual world paradigm (VWP; e.g., Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) is a methodology in which eye movements (relative to a display) are used to infer underlying language processing. This paradigm arguably allows for precise measurement (if relevant eye movements are closely time-locked to concurrent language) while using tasks that more closely reflect everyday language use (Tanenhaus et al., 1995). However, VWP displays (usually arrays of objects or images) differ from the everyday visual world (Henderson & Ferreira, 2004) in that (among other things) they limit object array size. Currently, little is known about how array size affects eye movements in the VWP. If changes in array size affect eye movement patterns in response to the same verbal stimulus, then the direct, time-locked link between eye movements and language may, in fact, be mediated by other cognitive processes.

EXPERIMENT 1

In order to examine the effects of array size on eye movements, we chose to replicate and extend a robust VWP finding. Huettig and Altmann (2005) found that a given noun elicited more immediate saccades to a target object (a referent of the noun) or a semantic competitor object (from the same

Please address correspondence to Karl G. D. Bailey, Behavioral Sciences Department, Andrews University, Berrien Springs, MI 49104–0030, USA. E-mail: kgbailey@andrews.edu

category as the noun), than to nonsemantically related distractors. These object images were displayed in 2×2 arrays. The current study examined whether the same patterns were present in 3×3 and 4×4 arrays. Two display types were selected from the Huettig and Altmann (2005) study. The target-present display contained both a target and competitor, along with distractors; the target-absent display contained only a competitor in addition to distractors. For each of 30 nouns, we constructed six displays, corresponding to the two display types and three array sizes. Each noun was recorded in a neutral utterance (material prior to the noun did not bias participants' eye movements).

Twelve participants listened to the 30 sentences while viewing accompanying displays on a computer screen. Each participant saw only one of the six display conditions for each sentence, but saw an equal number of each display condition overall. Participants were told to look wherever they wanted, as long as they were viewing the screen; no other task was given. Displays appeared one second prior to utterance onset; each display was visible for 8 s. Eye position data relative to the screen was sampled at 30 Hz. Analyses of immediate saccades to objects were conducted, relative to the noun, on a trial by trial basis.

Because the current VWP approach assumes that eye movement patterns are not affected by array size (sizes are chosen arbitrarily), we should not expect a main effect of array size. However, as in Huettig and Altmann (2005), we should expect to find more immediate saccades to the target or competitor than to distractors.

Our results replicated the expected semantic effects for the 2×2 array: participants made more saccades during the noun to the target (41.7% of trials) and competitor (26.7% of trials) in the target-present display than to individual distractors (17.5%). Likewise, in the target-absent display, participants made more saccades to the competitor (33.3%) than to individual distractors (17.5%). This pattern did not hold for the 3×3 and 4×4 arrays. As array size increased, differences between the target or competitor and individual distractors were reduced; in the 4×4 array, as many saccades were launched during the noun to the target or competitor as to individual distractors. However, the target and competitor were overall more likely to be fixated in the 3–4 s following the noun (until the offset of the display) than were distractors; thus, as array size increased, the expected overall patterns were obtained, but were less closely time-locked to language processing (Figure 1).

EXPERIMENT 2

The above array size effect may be due to participants' failure to fixate targets in the larger arrays prior to noun onset, thus precluding those objects
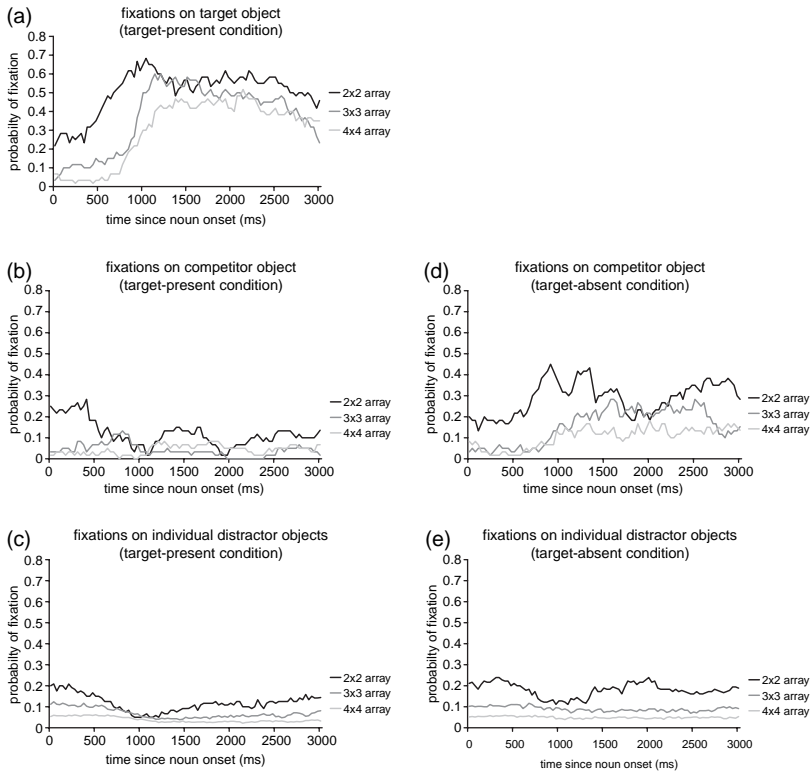
Figure 1. (a–e) These graphs show the probability of fixation on objects over time in Experiment 1, starting with critical noun onset. Note that additional objects in the 3 × 3 and 4 × 4 arrays lead to delayed responses to the critical noun and a weaker eye movement-language processing link in (a) and (d).

from being present in memory. Thus, in this experiment, we increased preview prior to utterance onset from 1 to 16 s for all three array sizes. Twelve participants took part in this experiment, which was otherwise identical to Experiment 1.

While the target was generally fixated during preview (even in the 4 × 4 array), neither target nor competitor effects were found for the 2 × 2 and 4 × 4 arrays. The 3 × 3 array elicited more immediate saccades to the target than to either the competitor or individual distractors. Again, participants fixated both targets and competitors more overall than distractors during the 3–4 s following the noun. It seems, then, that although targets were in long-term memory, some additional mediating process was necessary to drive eye movements.

EXPERIMENT 3

In order to control for array size, we ran a variation of Experiment 2 with 12 participants, in which 1 s of preview time was given per object (i.e., each array size was seen for a different length of time). Results were similar to Experiment 1, although competitor effects were not present in any target-present condition. Again, participants fixated the target and competitor more often overall relative to individual distractors. As in Experiment 2, this implicates a mediating factor in the language-eye movement link.

DISCUSSION

This study suggests that the VWP may be affected by the properties of displays. Specifically, increases in array size and preview time serve to decouple the strong, immediate link between language processing and eye movements. One possible mediator of the link may be short-term or working memory; larger arrays may result in the representation of the target object decaying or being dropped from short-term storage, thus requiring a visual search to take place. The lack of preview benefit is consistent with such an account. This study, then, suggests that in order to properly interpret VWP studies, models of cognition are needed that fully describe the link between language processing and eye movements.

REFERENCES

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107.

Henderson, J. M., & Ferreira, F. (2004). The interface of language, vision, and action: Eye movements and the visual world. New York: Psychology Press.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, *96*, B23–B32.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Dissociating mechanisms involved in accessing identity by dynamic and static cues

Leslie L. Steede

University of Sussex, Falmer, UK

Jeremy Tree

University of Exeter, Exeter, UK

Graham J. Hole

University of Sussex, Falmer, UK

Research indicates that idiosyncratic facial (review in Roark, Barrett, Spence, Abdi, & O'Toole, 2003) and bodily movements (Loula, Prasad, Harber, & Shiffrar, 2005) can provide useful cues for recovering person identity. However, research has not yet explored the cognitive or neural mechanisms that might be involved. Some researchers have speculated that the posterior superior temporal sulcus (pSTS) might be implicated in processing idiosyncratic facial and bodily movements (Roark et al., 2003). Others have suggested that that the same neural structure involved in static face recognition (the "fusiform face area" of the lateral fusiform gyrus) (Haxby, Hoffman, & Gobbini, 2000) might be involved in dynamic face recognition.

Here we aimed to provide insights into these issues, by exploring whether the mechanisms involved in accessing identity by dynamic facial and/or bodily cues, are likely to be the same or different to those involved in accessing identity by static facial cues. Across four experiments, two developmental prosopagnosics (CS and AA) as well as control groups, were tested. In Experiments 1 and 2, CS was tested on his ability to discriminate and learn to name identities by their idiosyncratic facial movements. In Experiments 4 and 5, AA completed two different tasks which tested his ability to learn to recognize identities by their idiosyncratic facial and bodily movements.

Both prosopagnosics have no known neural deficits, and perform in the normal range on general tests of memory and intellect. They are also unimpaired at tests of visual and object space perception, and memory tests for words and objects. However, both cases are impaired relative to matched control groups on tests of familiar face recognition devised in our lab, and

Please address correspondence to Mr. Leslie Steede, Department of Psychology, School of Life Sciences, Pevensey Building, University of Sussex, Falmer, BN1 9QH, UK. E-mail: lls21@sussex.ac.uk

standard tests of unfamiliar face memory (Duchaine & Nakayama, 2006; Warrington, 1984).

EXPERIMENTS 1–3

In Experiments 1–3, we used a variant of a task devised by Hill and Johnston (2001) to determine whether CS could discriminate between identities on the basis of their idiosyncratic facial movements. We used the same stimuli as Hill and Johnston[1] in which facial movements from live actors were projected onto the same 3-D head model. On each trial, three faces were presented sequentially. Two faces were animated with movements from the same individual, and one face was animated with movements from a different individual. Participants have to identify the "odd one out", or the different dynamic identity. In Experiment 1, faces were animated with rigid (rotational head movements) and nonrigid (expressive facial movements) movements. In Experiment 2, only rigid movements were presented. Across both experiments, CS's performance was on par with the performance of eight controls matched for age and level of education, and eight under-graduate control participants (CS's mean performance for Experiments 1 and 2 = 76% and 81%, respectively; control groups' mean performance for Experiments 1 and 2 = 71%; SD = 12% and 68%; SD = 12% respectively). In Experiment 3, 16 undergraduate participants completed a variant of Experiment 1, with the stimuli presented upside down. Face inversion leaves low-level motion cues intact, whilst disrupting an observer's ability to utilize face motion cues for face processing. Performance with inverted faces (M = 58%; SD = 7%) was significantly lower than performance with upright faces, $t(23.37) = 3.36$, $p < .0001$. This indicates that CS's performance was unlikely to reflect the use of low-level motion.

EXPERIMENTS 4–6

In Experiments 4–6, we tested whether AA could learn to recognize five identities on the basis of their idiosyncratic facial (Experiment 4) and bodily movements (Experiment 5). Both experiments consisted of three learning and test sequences. During the learning sequences, participants were required to associate a name with five idiosyncratic facial or bodily movements. During the test sequences, they were presented with a different example of the five dynamic identities that were learned, as well as two examples of dynamic identities that were not learned. Participants were

---

[1] We are very grateful to Harold Hill and Alan Johnston for allowing us to use this stimulus set.

required to provide a name only if they thought that one of the learned identities was presented. For Experiment 4, we selected different examples of the seven dynamic facial identities used in Experiment 1, for the learning and test sequences. For Experiment 5, we created a set of biological motion stimuli in which seven different actors were filmed in the same black tracksuit and footware as they jogged in a contained area. We occluded their heads during learning and test sequences (see Figure 1a). During the test sequences, a different video clip showing each actor jogging was thesholded and software was used to resize each bodily form to remove idiosyncratic static cues that could support recognition (see Figure 1a). Importantly, for both experiments, AA's performance was on par with six control participants matched for age and level of education (see Figure 1b). In addition, a one-way repeated measures ANOVA on the proportion of correct responses across time for Experiment 5 was significant, $F(2, 10) = 7.80$, $p = .008$, indicating that performance significantly increased across the three test times. However, the linear increase across time in Experiment 4 was only marginal ($p = .064$). In Experiment 6, six undergraduate participants were tested with the gait stimuli presented upside down. Performance with
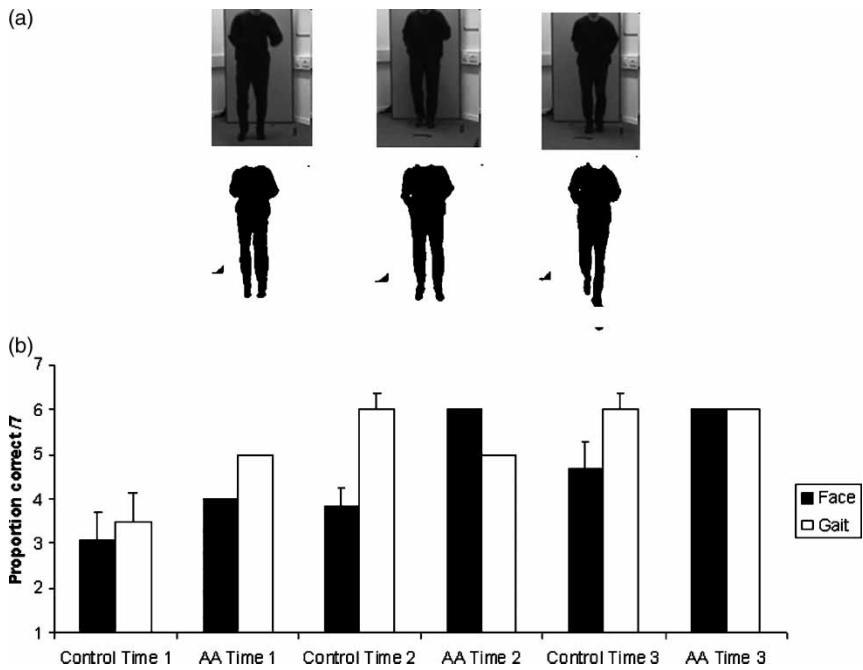

Figure 1. (a) Examples of the learning stimuli (top three bodies) and test stimuli (bottom three bodies) used for Experiment 5. (b) The proportion of correct responses for AA and controls for Experiment 4 (black bars) and Experiment 5 (white bars). Error bars show the standard error.

inverted bodily stimuli was significantly worst than when the stimuli were presented upright, $F(1, 10) = 37.24$, $p < .0001$. This indicates that AA's performance was unlikely to reflect low level motion processing.

Taken together, the results of these experiments suggest that the mechanisms involved in accessing identity by dynamic facial and bodily movements, are likely to be different to those involved in static face recognition. Future neural studies will be needed to establish whether the pSTS is involved in accessing identity by idiosyncratic movements. Recent behavioural research showing that observers are better able to recognize themselves, compared to their friends or strangers from idiosyncratic bodily movements (Loula et al., 2005), indicates that structures of the human mirror neuron system, which are believed to share motor and visual representations of familiar human actions, might play a role in accessing identity by idiosyncratic movements (review in Knoblish & Flach, 2003).

REFERENCES

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*, 576–585.

Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology*, *11*, 880–885.

Knoblich, G., & Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition*, *12*, 620–632.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 210–220.

Roark, D. A., Barrett, S. E., Spence, M. J., Abdi, H., & O'Toole, A. J. (2003). Psychological and neural perspectives on the role of motion in face recognition. *Behavioral and Cognitive Neuroscience Reviews*, *2*, 15–46.

Warrington, E. K. (1984). *Recognition memory test*. Windsor, UK: NFER-Nelson.

Is object-based attention mandatory?

Menahem Yeari and Morris Goldsmith

University of Haifa, Haifa, Israel

A central issue in visual attention research concerns the nature of the unit of attentional selection (e.g., Scholl, 2001). According to the *space-based* view,

Please address correspondence to Meni Yeari, Institute of Information Processing and Decision Making, University of Haifa, Haifa, Israel 31905. E-mail: myeari@gmail.com

attention selects unparsed regions of space in a manner analogous to a spotlight illuminating those regions. According to the competing *object-based* view, attention selects perceptual objects—products of preattentive segmentation and grouping processes operating on the basis of Gestalt principles and/or uniform connectedness. The space-based view suggests that attention is allocated irrespective of perceptual-organizational factors. In contrast, the object-based view posits mandatory organizational constraints on selection: the attempt to select an unparsed region of an object necessarily yields a processing advantage for all parts and features of the selected object (Kahneman & Henik, 1981).

The present study examined whether object-based attention is indeed mandatory, or rather, under strategic control: under conditions in which it is strategically worthwhile to allocate attention in a space-based manner, will attention nonetheless be affected by organizational factors, or will perceivers be able to resist the influence of such factors? Surprisingly, despite a large amount of related research, there is still no clear answer to this question.

The claim that object-based attention is mandatory has been supported implicitly by the many studies in which object-based effects were observed even though object-based allocation was neither strategically expedient nor explicitly required by the task (e.g., Egly, Driver, & Rafal, 1994). Note, however, that although such findings might reflect a mandatory mode of allocation which is impervious to strategic control, they might rather simply reflect a general tendency or *default* mode that is used unless there is some special (e.g., strategic) reason to do otherwise.

Evidence against the claim of mandatory object-based attention is similarly inconclusive. It is now well established that object-based effects on attention are not observed under all conditions (e.g., Goldsmith & Yeari, 2003), implying that object-based selection is not *universal*. It is still quite possible, however, that object-based selection is *mandatory* whenever the "boundary conditions" for such selection are met. In order to refute the claim that object-based attention is mandatory (involuntary), one must show that object-based selection can be avoided when it is strategically worthwhile to do so, under conditions in which it would otherwise be observed. Two studies that directly manipulated the expediency of object-based versus space-based attention, by manipulating the probabilities of same-object/group or different-object/group targets, found that grouping effects were attenuated but not eliminated in a serial scanning task (Beck & Palmer, 2002) and that object-based effects were eliminated at long but not short cue–target SOAs in a spatial cueing task (Shomstein & Yantis, 2004). Thus, until now, no study has shown that the effects of perceptual organization on attention can be completely avoided through strategic control.

METHODS AND RESULTS

In the current study we used a very strong manipulation of expediency to induce participants to avoid an object-based mode of attentional allocation. In two experiments using a specially designed spatial-cueing discrimination task, participants fixated attention on a central cue, which was part of a perceptual group (Experiment 1; Figure 1A) or uniformly connected object
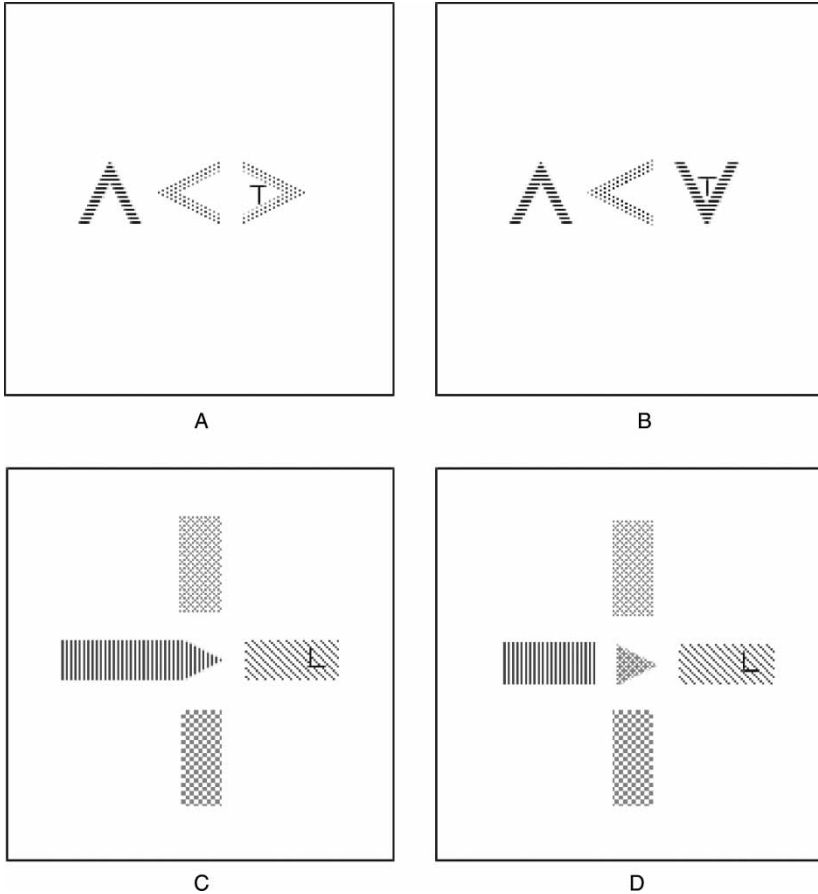

Figure 1.   The visual displays used in the experiments. A central arrow-shape acts as a spatial cue, and the peripheral shapes mark the potential target locations (different stimulus textures in the figure represent different stimulus colours used in the actual experiments): (A) Experiment 1—grouping between cue and peripheral location marker. (B) Control—no grouping. (C) Experiment 2—cue and peripheral location marker form a uniformly connected object. (D) Control—cue and marker do not form a uniformly connected object. The target is a letter "L" or "T" (uncued same-object target in panel A; cued different-object target in panel C).

(Experiment 2; Figure 1C). The cue always pointed to an opposite, different-object location. By varying cue validity, the strategic incentive to avoid preferential attention to the entire cue-object was manipulated: when the cue was undiagnostic of target location, such that same-object (as the cue) and different-object targets were equally likely (no/weak incentive to avoid object-based attention), a significant ($\alpha = .05$) same-object RT advantage was observed ($M = 11$ ms in Experiment 1; 20 ms in Experiment 2) across all levels of cue–target SOA (100–400 ms in Experiment 1; 100–300 ms in Experiment 2). However, when the cue was highly valid, such that cued different-object targets were much more likely than uncued same-object targets (strong incentive to avoid object-based attention), the same-object RT advantage disappeared. Instead, spatially cued different-object RTs were significantly faster than same-object RTs at SOA ≥ 200 ms ($M = 14$ ms in Experiment 1; 12 ms in Experiment 2), with no significant difference between uncued different-object RT and uncued same-object RT (Experiment 2, all Fs < 1) at any SOA.

Moreover, to examine whether there was any residual cost of resisting object-based attention, we compared performance in the valid cueing condition with performance in a control condition in which the central cue was not grouped (Figure 1B) or connected (Figure 1D) to any of the peripheral location markers. Performance was no different when cue–object grouping or connectedness had to be overcome than when there was no such grouping or connectedness at all (all Fs < 1 for both experiments). Overall, these results suggest that object-based attention is a default mode of allocation which, at least under some conditions, can be completely overridden by strategic control.

DISCUSSION

In light of these findings, we propose a parsimonious way of reconciling object-based and space-based attention: Attention tends to be allocated to the most relevant-informative unit of space, taking into account both enduring and transitory strategic considerations. From an evolutionary perspective, it is generally most expedient to allocate attention to entire objects, because of their general ("default") informational advantage over unorganized (meaningless) stimuli falling in arbitrary spatial regions (Yantis, 1996). However, attention can also be allocated in a manner that disregards object and group boundaries, if this is called for (strongly enough) by strategic considerations in the task at hand.

REFERENCES

Beck, D. M., & Palmer, S. E. (2002). Top-down influences on perceptual grouping. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1071–1084.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*, 161–177.

Goldsmith, M., & Yeari, M. (2003). Modulation of object-based attention by spatial focus under endogenous and exogenous orienting. *Journal of Experimental Psychology: Human Perception and Performance*. *29*, 897–918.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*, 1–46.

Shomstein, S., & Yantis, S. (2004). Configural and contextual prioritization in object-based attention. *Psychonomic Bulletin and Review*, *11*, 247–253.

Yantis, S. (1996). Attentional capture in vision. In A. F. Kramer & M. G. H. Coles (Eds.), *Converging operations in the study of visual selective attention*. Washington, DC: American Psychological Association.