

Short article

Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements

James R. Brockmole

University of Edinburgh, Edinburgh, UK

John M. Henderson

Michigan State University, East Lansing, MI, USA

When confronted with a previously encountered scene, what information is used to guide search to a known target? We contrasted the role of a scene's basic-level category membership with its specific arrangement of visual properties. Observers were repeatedly shown photographs of scenes that contained consistently but arbitrarily located targets, allowing target positions to be associated with scene content. Learned scenes were then unexpectedly mirror reversed, spatially translating visual features as well as the target across the display while preserving the scene's identity and concept. Mirror reversals produced a cost as the eyes initially moved toward the position in the display in which the target had previously appeared. The cost was not complete, however; when initial search failed, the eyes were quickly directed to the target's new position. These results suggest that in real-world scenes, shifts of attention are initially based on scene identity, and subsequent shifts are guided by more detailed information regarding scene and object layout.

The efficient detection of goal-relevant information can be guided by memory for stable visual contexts that in turn predict the location of task-relevant visual information, an effect called *contextual cueing* (see Chun, 2003, for a review). In contextual cueing paradigms, repeated exposure to a specific arrangement of target and

distractor items leads to progressively faster search for the target. Common search displays used in these demonstrations include randomly arranged letters (Chun & Jiang, 1998), novel 2-D shapes (e.g., Chun & Jiang, 1999), or 3-D volumetric shapes (e.g., Chua & Chun, 2003). Strikingly, these learning effects occur despite

Correspondence should be addressed to James R. Brockmole, Psychology Department, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK. Email: James.Brockmole@ed.ac.uk

This research was supported by the National Science Foundation (BCS-0094433 and IGERT grant ECS-9874541), the US Army Research Office (W911NF-04-1-0078), and a Strategic Partnership Grant from the MSU Foundation to JMH. We thank Sonia Dutt and Paul Turnpaugh for their help with data collection.

We thank Alexander Pollatsek and Mark Lansdale and an anonymous reviewer for their helpful comments on an earlier draft of this article.

observers' inability to explicitly discriminate repeated displays from novel stimulus arrangements at the conclusion of the experiment.

In order for search times through a repeated visual display to decrease as a function of learning, two processes must be successfully engaged: recognition and attention guidance. The display must be recognized (in this case implicitly) in the sense that it must be matched to a corresponding representation stored in memory. Following recognition, the encoded pairing between the display and target location must guide attention to the search target. Peterson and Kramer (2001) recorded observers' eye movements through repeated sets of letter arrays to examine how each of these processes proceeds during search. They reasoned that, given the coupling between the locus of attention and the locus of fixation, once a repeated display is recognized, the eyes should be biased to move toward the target. The proximity of each fixation to the target was taken as a measure of the precision of the guidance system. Over the course of learning, recognition of the repeated displays was rarely immediate. Even in the final epoch of learning (the 12th through 16th repetition of each old display), fewer than 15% of the first observer-selected fixations fell on the target. For the remaining trials, both the Euclidean and angular distance between the first observer-selected fixation point and the target did not differ between repeated and novel displays. However, in these cases where identification of the arrays was delayed, and no bias for the initial eye movement to be directed to the target was observed, fewer fixations were ultimately required to locate the target than for novel displays. Thus, although recognition of repeated contexts sometimes occurred immediately, on the majority of trials recognition occurred after search had commenced. This finding accounts for the failure of search slopes in contextual cueing experiments to fall to 0 ms/item, even over the course of extensive training (e.g., Chun & Jiang, 1998).

Because research on contextual cueing has mostly been limited to arrays of shapes and letters, little is known about how repeated experience with real-world scenes affects the guidance of

attention to known target positions. Indeed, a recent report argues that differences between contextual cueing in stimulus arrays and real-world scenes exist (Brockmole & Henderson, 2006; see also Brockmole & Henderson, *in press*). When observers were repeatedly exposed to photographs of real-world scenes each containing a consistently but arbitrarily positioned target letter, repetition led to a decrease in search time. However, scene-target associations were learned up to 5 times faster and led to a search benefit 20 times greater than the cueing effects observed in non-scene displays. These additional benefits were driven by at least two mechanisms not available to support the learning of context-target associations in nonscene stimuli.

First, scene-target associations were explicitly encoded in memory. Observers recognized repeated scenes and recalled their associated target positions far better than they did for novel scenes. This result is consistent with the finding that people are able to explicitly recognize hundreds if not thousands of previously novel scenes after a single exposure (Shepard, 1967; Standing, 1973). Second, cueing was facilitated by semantic memory for scene content. Approximately twice the number of repetitions were required to observe a maximal learning benefit when scenes were inverted (making them harder to interpret) compared to upright scenes. This finding is consistent with research demonstrating that scenes can be quickly (<100 ms) categorized and identified, engendering expectations about layout and component objects (Friedman, 1979; Potter, 1976; see also Henderson & Ferreira, 2004), which in turn can guide attention to informative scene regions determined by one's task (Henderson, Weeks, & Hollingworth, 1999; Loftus & Mackworth, 1978; Mackworth & Morandi, 1967; Yarbus, 1967; see Henderson, 2003).

The present report marries contextual cueing studies using real-world scenes (Brockmole & Henderson, 2006) and the use of eye movements to study contextual cueing (Peterson & Kramer, 2001) to examine how scene recognition and attention guidance unfold very early in real-world scene viewing. Given the speed with

which real-world scenes are identified and recognized, we expect attention, and the eyes, to proceed to the target faster than has been previously reported for arrays of letters. However, what is not known is the nature of the information used to guide attention to known target positions. Scenes can be represented in terms of their identity or basic-level concept or in terms of their specific arrangements of visual properties. Are both types of representation available in equal time, or is one type of information available to guide the eyes sooner than the other?

To address this question, the experiment was divided into two phases. In the first phase, observers were repeatedly shown photographs of real-world scenes that contained consistently

but arbitrarily located targets (Brockmole & Henderson, 2006). This *learning phase* allowed observers to associate the target position with the context provided by the scene and enabled search to become more efficient. In the *transfer phase*, learned scenes were unexpectedly mirror reversed, spatially displacing local objects and features while preserving the scene's identity (see Figure 1). As a result, the target's position shifted relative to the observer's egocentric reference frame. The question of interest was whether and how prior contextual cueing would transfer to these altered stimuli.

If attention is initially directed by recognition of the identity and/or basic-level category of the scene without regard to the specific arrangement of visual features, the mirror reversal would not



Figure 1. Example stimuli that were repeated during the learning and transfer phases of the experiment (stimuli were presented in full colour). The transfer stimuli are mirror reversals of the learning stimuli. Target size and position are highlighted with white squares. In the first example, the target is located on the side panel of the moped. In the second example, the target is located on the trunk of the large tree.

be noticed, and search for the target in the transfer phase should commence in the same manner as it did at the conclusion of the learning phase. That is, attention would move in the direction of the original position of the target in the display. Over time, as spatial structure and detail is processed during the course of viewing, subsequent recognition of the mirror reversal may occur, and attention would then be guided to the target's new position. In this case, the benefits of contextual cueing would be reduced when the scenes are mirror reversed because attention is shifted in the wrong direction, but would not completely dissipate as later identification of the mirror reversal occurs. This outcome is predicted by previous work that has argued that scenes are initially processed in terms of their identity or basic-level category membership with more detailed information about objects and their locations collected over the course of viewing (Oliva, 2005).

On the other hand, if attention is primarily directed by the recognition of specific visual features, then attention should not initially move toward the target's old position. Under this account, two possible outcomes exist. First, mirror reversals could be inconsequential to recognition if the realignment of features can be immediately noticed. This would result in initial movements of attention toward the new correct location. That is, learned context-target associations would transfer to the mirror-reversed stimuli with no adverse effects on search. This outcome is supported by previous work demonstrating that observers are very good at discriminating previously viewed scenes from mirror-reversed lures (Standing, Conezio, & Haber, 1970). Second, it is possible that mirror reversals may lead to a total failure in recognizing a repeated scene, as a direct matching of features present in the mirror-reversed image would not match the encoded representation in memory. This would result in random initial shifts of attention, and so search in the transfer phase would resemble that observed for the novel trials.

Although search times can reflect disruptions in search, they do not pinpoint the cause of that disruption. Therefore, we recorded observers' eye

movements during search. If mirror reversals disrupt search, then the number of fixations needed to find the target should increase, and a decrease in bias toward the target should be observed relative to the learning phase. To quantify any bias for the eyes to move toward the target, we calculated a *scan pattern ratio* (Henderson, Weeks, & Hollingworth, 1999). This ratio contrasts the total distance traversed by the eyes through the scene (the summed spatial displacement of all saccades) and the minimum distance needed to move in a straight line from the initial fixation point to the target. The more direct the route the eyes travel between the initial fixation point and the target, the lower the scan pattern ratio. In addition, in the transfer phase, eye movement direction was also quantified by assessing whether each subsequent eye movement moved closer to the original target position or to the new target position.

Method

A total of 8 Michigan State University undergraduates participated after providing informed consent and were compensated with course credit. Observers searched for a known target (a small grey "T" or "L") embedded within 144 full-colour photographs of real-world scenes (Figure 1). Each photograph was displayed at a resolution of 800 by 600 pixels by 24-bit colour on a 19-inch cathode-ray tube (CRT) with a refresh rate of 100 Hz. Although the stimuli were the same as those used by Brockmole and Henderson (2006), the present report used an independent sample of observers and different apparatus. Chin and forehead rests maintained a viewing distance of 57 cm. Under these conditions, the photographs subtended 37 deg horizontally and 27.5 deg vertically, and the targets subtended 0.25 deg horizontally and vertically. Eye position was monitored using an ISCAN pupil and corneal reflection tracking system sampling at 240 Hz accurate to within 0.5 deg of visual angle.

The sequence of trials was divided into 17 blocks of 16 trials. Blocks 1–10 constituted the

learning blocks, which randomly intermixed 8 *novel trials* with 8 *repeated trials*. A novel trial presented a scene that had not been previously shown in the experiment. A repeated trial presented one of eight scenes that was previously shown. Critically, the target's location in each repeated scene was fixed, although the target's identity was randomly selected with each repetition. Blocks 11–17 randomly intermixed 8 novel trials with 8 *transfer trials*. A transfer trial presented a mirror-reversed version of a repeated scene. We refer to Block 11 as the *transfer block* and Blocks 12–17 as *relearning blocks*. Under these constraints, the order of trials was randomly selected for each participant. No information regarding the block structure or the repetition of scenes was given to observers.

Participants began the experimental session by completing a calibration routine that served to map the output of the eyetracker onto display position. Calibration was monitored throughout the experiment and was adjusted when necessary. At the beginning of each trial, a blue dot was centred on a grey background. Observers were instructed to look at this dot and to press a key when ready to view the scene. Upon identifying the target, observers pressed one of two buttons corresponding to either "T" or "L". The trial was

terminated if a response was not made within 20 s of scene onset. Observers were told to respond as quickly but as accurately as possible.

Results

Analyses excluded trials on which no response was given within the allotted 20-s viewing window (5% of trials) or if a response was incorrect (2% of trials).

Number of fixations to target

Number of fixations was defined as the number of discrete fixations on a scene before the target was fixated, including the first experimenter-induced fixation. For all analyses, the target was considered to be fixated if the point of fixation was within 40 pixels of the centre of the target. Trends in the number of fixations to the target are illustrated in Figure 2.

Learning blocks. A repeated measures analysis of variance (ANOVA) demonstrated main effects of trial type (novel vs. repeated), $F(1, 7) = 16.6$, $p < .01$, and block, $F(9, 63) = 2.27$, $p < .05$. Critically, these factors interacted, $F(9, 63) = 6.49$, $p < .001$. Considering novel trials only, no

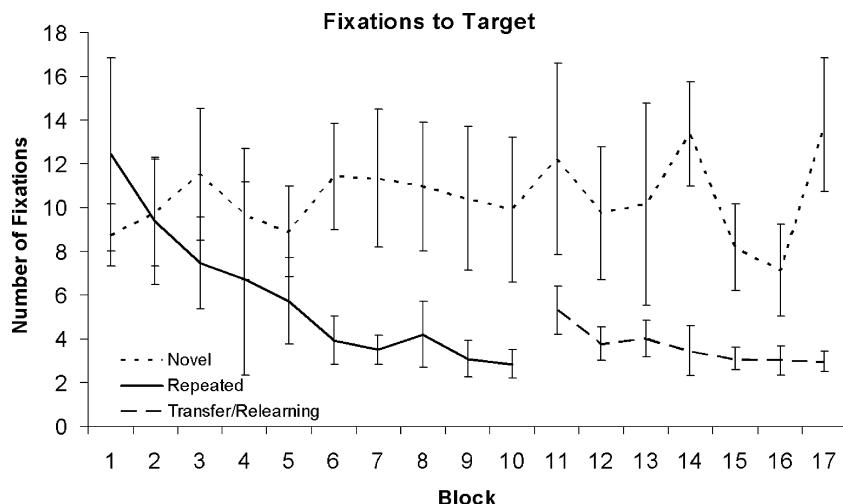


Figure 2. Average number of fixations required to find the target on novel and repeated stimuli in the learning, transfer, and relearning blocks. Error bars represent 95% confidence intervals.

effect of block was observed, $F < 1$. On average, 10.2 fixations intervened between the start of the trial and the first fixation on the target. The number of fixations prior to fixating the target in repeated trials, however, decreased over block, $F(9, 63) = 10.2$, $p < .001$, from an average of 12.4 in Block 1 to 2.8 in Block 10.

Transfer Block 11. On average, 5.3 fixations were needed to locate the target on the transfer trials, a reliable increase compared to the repeated trials in Block 10, $t(7) = 5.63$, $p < .001$. However, fewer fixations to the target were required than for the novel trials in the same block, which averaged 12.6 fixations, $t(7) = 3.22$, $p < .05$. Thus, the initial mirror reversal led to a search cost, but some transfer of learning to the mirror-reversed stimuli occurred.

Relearning blocks. A repeated measures ANOVA demonstrated main effects of trial type, $F(1, 7) = 182$, $p < .001$, and block, $F(6, 42) = 3.30$, $p < .01$. Critically, these factors interacted, $F(6, 42) = 2.27$, $p < .05$. Considering novel trials only, although an effect of block was observed, $F(6, 42) = 2.57$, $p < .05$, no interpretable trends were present. On average, 10.6 fixations intervened between the start of the trial and the first fixation on the target. For repeated trials, however, the number of fixations to the target decreased with block, $F(6, 42) = 9.39$, $p < .001$, from an average of 5.3 on Block 11 to 2.9 on Block 17, a level of performance that did not reliably differ from that observed at the end of the learning blocks (Block 10), $t(7) < 1$.

Scan pattern ratio

The scan pattern ratio is a measure of how directly the eyes move to the target (see Henderson et al., 1999). The scan pattern ratio divides the total distance travelled by all eye movements prior to arriving at the target by the linear distance between the initial fixation point and the target:

$$\frac{\sum_{i=1}^n S_i}{\sqrt{(x_T - x_F)^2 + (y_T - y_F)^2}} \quad (1)$$

where S_i is the spatial distance subtended by an individual eye movement, n is the number of fixations to the target, (x_T, y_T) is the position of the target, and (x_F, y_F) is the position of the first fixation point. As the eye movement pattern becomes less direct, the scan pattern ratio increases. Trends in scan pattern ratio are illustrated in Figure 3.

Learning blocks. A repeated measures ANOVA demonstrated main effects of trial type, $F(1, 7) = 13.5$, $p < .01$, but not block, $F(9, 63) = 1.41$. Critically, however, these factors interacted, $F(9, 63) = 3.22$, $p < .01$. Considering novel trials only, no effect of block was observed, $F(9, 63) < 1$. On average, the scan pattern ratio was 8.0. The scan pattern ratio in repeated trials, however, decreased over blocks, $F(9, 63) = 7.75$, $p < .001$, from an average of 8.8 on Block 1 to 2.0 on Block 10.

Transfer block. The initial mirror reversal led to a search cost. On average, the scan pattern ratio on transfer trials was 4.3, a reliable increase compared to Block 10, $t(7) = 5.05$, $p < .001$. However, the scan pattern ratio on the transfer trials was less than half that on the novel trials in the same block, which averaged 9.2, $t(7) = 2.72$, $p < .05$, again indicating that some transfer of learning to the mirror-reversed stimuli occurred. These results again indicate that the initial mirror reversal led to a search cost, but that some transfer of learning to the mirror-reversed stimuli occurred.

Relearning blocks. A repeated measures ANOVA demonstrated main effects of trial type, $F(1, 7) = 177$, $p < .001$, and block, $F(6, 42) = 2.78$, $p < .05$. Critically, these factors interacted, $F(6, 42) = 3.26$, $p < .01$. Considering novel trials only, a reliable effect of block was observed on novel trials, $F(6, 42) = 2.64$, $p < .05$, but no interpretable trends were reliable. On average, the scan pattern ratio was 8.6. For repeated trials, however, the number of fixations to the target decreased with block, $F(6, 42) = 9.88$, $p < .001$, from 4.3 on Block 11 to 2.2 on Block 17, a level of performance that did not reliably

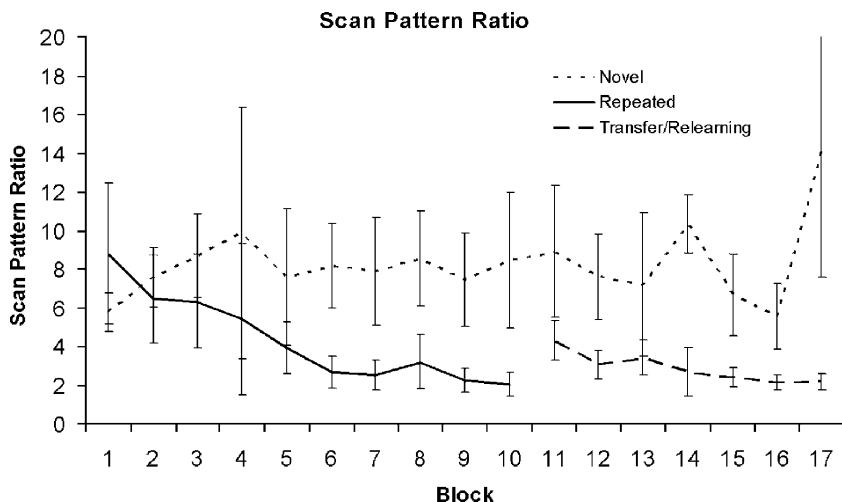


Figure 3. Average scan pattern ratio observed during search through novel and repeated stimuli in the learning, transfer, and relearning blocks. Error bars represent 95% confidence intervals.

differ from that observed at the end of the initial learning blocks (Block 10), $t(7) < 1$.

Eye direction

The previous analyses demonstrated that search was adversely affected when learned scenes were mirror reversed, although search behaviours did not return to prelearning levels. This pattern could be generated by a mixed distribution of trials or subjects on which performance is not disrupted and others on which performance is completely disrupted. Another possibility is that attention moves to the target's old position as the scene but not the mirror reversal is recognized, but when the target is not found the ensuing search is not "memoryless". That is, attention may initially move to the target's old position based on an initial recognition of identity or concept, but when the target cannot be found memory representations containing information beyond this identity (e.g., feature information) may guide attention to the target's new position via a relatively direct path.

To contrast these possibilities, we determined whether on the transfer trials the eyes initially moved toward the target's old position or if they instead moved toward the target's new position.

To quantify the direction of the eye movements, the linear distance between the locus of the eyes and the target's old and new positions was calculated for every fixation. The distance to the target's old position was subtracted from the distance to the target's new position. Thus, movements toward the old position yielded positive values while movements toward the new position produced negative values.

Average trends in eye direction on transfer trials in Blocks 11 and 17 are illustrated in Figure 4 broken down by fixation number. Figure 5 plots trends in eye direction for each individual participant. Fixation 1 is the first experimenter-induced fixation on the scene and was equidistant from the old and new target positions. A repeated measures ANOVA demonstrated main effects of block, $F(1, 7) = 14.6$, $p < .01$, and fixation number, $F(4, 28) = 19.3$, $p < .001$. Critically, these factors interacted, $F(4, 28) = 10.4$, $p < .001$. Single degree of freedom polynomial tests showed that in Block 11 the change in eye direction over fixations was characterized by a reliable quadratic trend, $F(1, 7) = 24.9$, $p < .01$, which first deflected positively (toward the original target position) before deflecting negatively (toward the new target position). For Block 17,

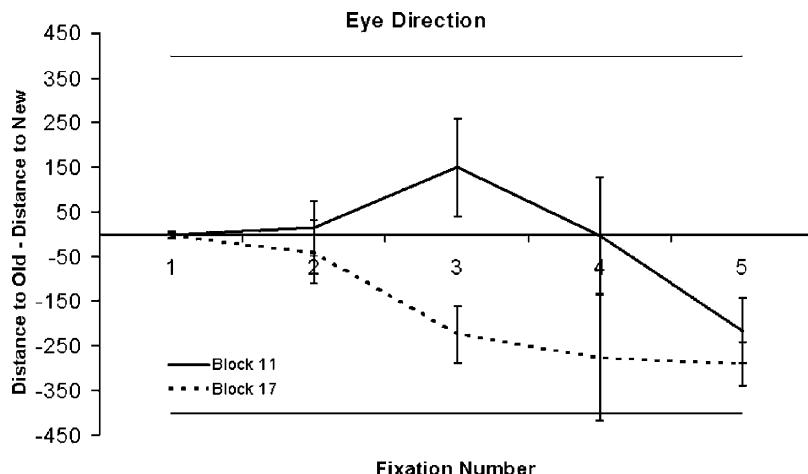


Figure 4. Average eye direction for the first five fixations from the beginning of search on Blocks 11 (transfer block) and 17 (end of relearning). Positive values indicate that fixation is closer to the target's old position; negative values indicate that fixation is closer to the target's new position. Error bars represent 95% confidence intervals. The horizontal lines represent the average distance between the target's old and new positions.

however, only a strictly decreasing linear trend was observed, $F(1, 7) = 45.3, p < .001$, showing that the eyes moved directly toward the target's new position. These results indicate that upon mirror reversal, initial eye movements were directed toward the target's old position, but subsequent eye movements were directed toward the target's new position.

Summary of dependent measures

Over the course of the learning blocks, a gradual decrease in the number of fixations needed to find the target as well as a decrease in scan pattern ratio were observed. When the repeated scenes were mirror reversed in the transfer block, the number of fixations needed to find the target and the scan pattern ratio increased, although not to levels observed for novel (unlearned) stimuli. Additionally, a propensity for the eyes to move toward the target's original position before moving to its new position was observed in the transfer block. By the end of the relearning blocks, the disruption to search caused by mirror reversals had dissipated as the eyes moved directly to the target's new position.

Discussion

The present study examined the recognition and attention guidance mechanisms involved in visual search when target location systematically covaries with its surrounding context, in this case a real-world scene. Previous contextual cueing studies using nonscene stimuli have indicated that recognition of repeated displays develops over the course of search, even following extensive training. Research in scene perception, however, has demonstrated that scenes can be identified well within the first 100 ms of viewing, indicating that scene recognition in contextual cueing paradigms may be much faster when real-world scenes serve as context. Indeed, the ease with which scenes can be identified influences the overall speed with which scene–context covariations are learned (Brockmole & Henderson, 2006). In the present report, after just 10 repetitions, the target was fixated within just 1.8 eye movements, indicating very fast scene recognition and highly accurate attentional guidance.

We also investigated whether scene recognition occurs in at least partially independent stages characterized by different time courses.

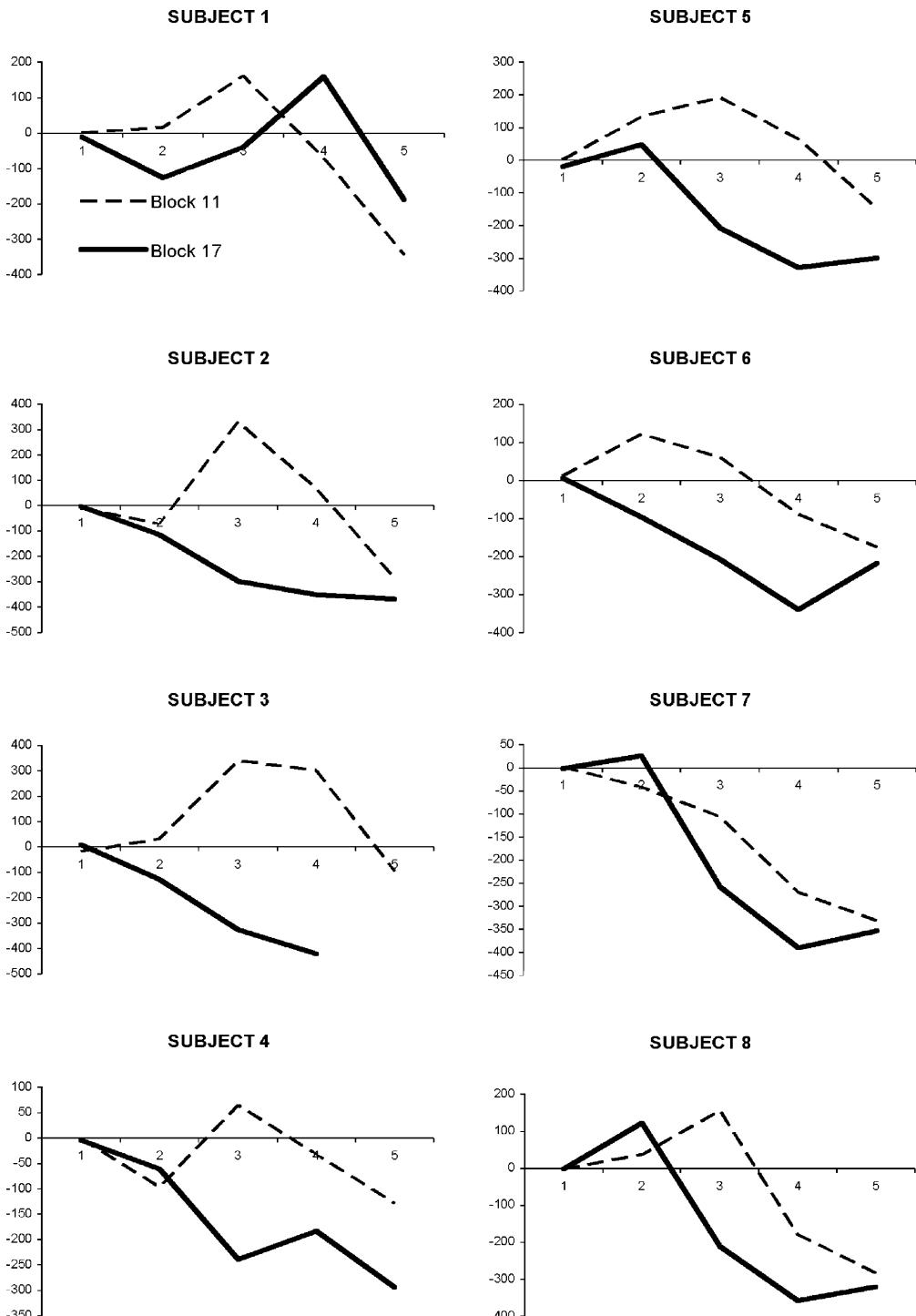


Figure 5. Average eye direction for the first five fixations from the beginning of search on Blocks 11 and 17, broken down by observer.

Specifically, scenes can be described in terms of their identities or basic-level concepts or in terms of their visual features and properties. To determine whether recognition of both types of information occur in unison, spatial configuration was manipulated independently of scene identity by mirror reversing learned scenes. In this situation visual information was displaced, but the scene's identity, category membership, and content were preserved. When first presented with a mirror reversal of a learned scene, observers initially moved their eyes toward the position in the display in which the target had previously appeared as if the mirror reversal was not noticed by the observers, a surprising result given that observers can easily discriminate studied scenes from mirror-reversed lures after a single exposure (Standing et al., 1970). This error in target localization caused an increase in search time¹ as additional fixations were required to locate the target. However, the disruption was not absolute; upon discovering that the target was not in the expected location, the eyes quickly moved toward the new target's position. Thus, although search was slowed, savings were observed compared to the novel scenes.

The sudden reorientation of attention toward the target's new position following an errant initial search is only possible if memory for more than scene identity was available. If participants simply learned an association between scene identity and location or direction from initial fixation, an extensive serial search would have ensued when the target was not at its expected location. Thus, the evidence suggests that contextual cueing operates in real-world scenes by at least two recognition processes. First, the identity of the scene is recognized without reference to the specific arrangement of visual features. This association initially guides attention to the target, perhaps by invoking a learned eye movement vector to the region of the display that once contained the

target. Second, other visual information in the scene is recognized, such as the reorientation of visual features relative to the observer. There are several possibilities for what the additional information might be. For example, it could be local feature information, local identity information, or global orientation information in which the mirror reversal is noticed, and a mental transformation is undertaken to locate the target. Following an initial shift of attention based on scene identity, subsequent shifts are guided by this additional source of information. Interestingly, though, such additional information does not appear to be used unless recognition of scene identity alone fails.

Original manuscript received 23 November 2005
Accepted revision received 21 February 2006

REFERENCES

- Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition*, 13, 99–108.
- Brockmole, J. R., & Henderson, J. M. (in press). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Chua, K.-P., & Chun, M. M. (2003). Implicit scene learning is viewpoint-dependent. *Perception & Psychophysics*, 65, 72–80.
- Chun, M. M. (2003). Scene perception and memory. *Psychology of Learning and Motivation*, 42, 79–108.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36, 28–71.
- Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, 10, 360–365.
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory

¹For archival purposes we note that average search time for novel scenes did not vary over blocks and averaged 4102 ms. For repeated scenes, average search time decreased from 4523 ms in Block 1 to 1329 ms in Block 10. In the transfer block, search time increased by an average of 792 ms. At the conclusion of the relearning blocks, search time averaged 1371 ms. All comparisons were statistically reliable.

- for gist. *Journal of Experimental Psychology: General*, 108, 316–355.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7, 498–504.
- Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.
- Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 210–228.
- Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 565–572.
- Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics*, 2, 547–552.
- Oliva, A. (2005). Gist of the scene. In L. Itti, G. Reese, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–256). San Diego, CA: Elsevier.
- Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics*, 63, 1239–1249.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2, 509–522.
- Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 6, 156–163.
- Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, 25, 207–222.
- Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 19, 73–74.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.