

The role of color in visual search in real-world scenes: Evidence from contextual cuing

KRISTA A. EHINGER

Massachusetts Institute of Technology, Cambridge, Massachusetts

AND

JAMES R. BROCKMOLE

University of Edinburgh, Edinburgh, Scotland

Because the importance of color in visual tasks such as object identification and scene memory has been debated, we sought to determine whether color is used to guide visual search in contextual cuing with real-world scenes. In Experiment 1, participants searched for targets in repeated scenes that were shown in one of three conditions: natural colors, unnatural colors that remained consistent across repetitions, and unnatural colors that changed on every repetition. We found that the pattern of learning was the same in all three conditions. In Experiment 2, we did a transfer test in which the repeating scenes were shown in consistent colors that suddenly changed on the last block of the experiment. The color change had no effect on search times, relative to a condition in which the colors did not change. In Experiments 3 and 4, we replicated Experiments 1 and 2, using scenes from a color-diagnostic category of scenes, and obtained similar results. We conclude that color is not used to guide visual search in real-world contextual cuing, a finding that constrains the role of color in scene identification and recognition processes.

Contextual cuing is a process by which long-term memory representations for previously encountered visual displays are used to facilitate search. When people are repeatedly exposed to a specific arrangement of target and distractor items, search becomes progressively faster (e.g., Chun & Jiang, 1998) and is driven by fewer and more direct eye movements toward the target (Brockmole & Henderson, 2006a; Peterson & Kramer, 2001). Thus, through experience with the visual world, people are able to learn the likely locations of particular objects in particular contexts and to use their memory for these associations to guide attention in visually based tasks (for reviews, see Chun, 2000, 2003; Jiang, Makovski, & Shim, in press; for an alternative view, see Kunar, Flusberg, Horowitz, & Wolfe, 2007). Demonstrations of contextual cuing have been obtained with arrays of randomly arranged letters (e.g., Chun & Jiang, 1998) and novel shapes (e.g., Chua & Chun, 2003; Chun & Jiang, 1999) and with naturalistic scenes (Brockmole, Castelano, & Henderson, 2006; Brockmole & Henderson, 2006a, 2006b). However, the speed and magnitude of the learning benefit varies across stimulus types, with natural-scene–target covariation being learned up to 5 times faster and yielding a benefit of up to 20 times greater.

The differences in learning speed and the magnitude of the search benefit across stimuli hint that different mechanisms may underlie contextual cuing, depending on the nature of the context available. Indeed, evidence suggests that whereas contextual cuing in arrays of letters and shapes is an

implicit-learning process (Chun & Jiang, 1998), learning is explicit when scenes are employed as context (Brockmole & Henderson, 2006b). Given this difference, at a minimum, it is clear that paradigms using nonscene stimuli have led only to a partial characterization of contextual cuing within the real world. Given the ubiquity of visual search in our everyday interactions with the visual world, it is therefore imperative that the mechanisms underlying contextual cuing in real-world scenes receive as much attention as has been prescribed to artificial stimulus arrays.

One critical issue concerning contextual cuing specifically, and the development of behaviorally functional scene memory generally, is the nature of the information that is used to reference target position. Although critical properties of visual displays have been extensively studied in simple arrays of letters and shapes, there have been few studies in which the question of which visual aspects of real-world scenes are critical to learning context–target contingencies has been considered. One recent exception contrasted the roles of global and local contexts for contextual cuing in naturalistic scenes. Whereas contextual cuing in simple stimulus arrays is driven by associations between the target and a small subset of nearby items (Jiang & Wagner, 2004; Olson & Chun, 2002), contextual cuing in scenes depends on stable covariation between the target and global scene features (Brockmole et al., 2006).

In this report, we focus on the relative importance of spatial configuration and object features in contextual

J. R. Brockmole, james.brockmole@ed.ac.uk

cuing. When identical displays are repeatedly presented to observers, it is not clear whether contextual cuing depends on memory for the spatial arrangement of objects, independently of the features that define those objects, or whether object features are also used to reference the target's location. To paraphrase a question posed by Jiang and Song (2005), as a man repeatedly walks along the streets of Boston or Edinburgh, does he extract the spatial layout of the city as if each landmark is a shapeless and colorless mass, or does he integrate the specific visual features of each landmark with the layout? Previous research has attempted to contrast these possibilities by constructing visual search tasks in which objects' visual features (e.g., shape or color) and locations are independently manipulated.

A common feature manipulation in past contextual-cuing experiments has been color (e.g., Jiang & Chun, 2001; Jiang & Song, 2005; Kunar, Flusberg, & Wolfe, 2006; Olson & Chun, 2002). Such experiments have indicated that the importance of color in the learning of context–target associations depends on the extent to which those features are relevant to search. For example, in situations in which color was an irrelevant aspect of search, it played no role in contextual cuing (Olson & Chun, 2002). When items were either red or green and one of these color sets predicted target position, contextual cuing was equivalent whether or not the target shared the same color as the predictive distractors, suggesting that spatial features play a more dominant role than surface features in contextual cuing. However, spatial context learning can be contingent on object features when those features are predictive of target location. For example, Jiang and Song showed observers displays, some of which contained black items and some of which contained white items. A subset of these displays was repeatedly presented and, as was expected, learning benefits were observed as search times decreased across repetitions. However, this learning benefit was subsequently eliminated when new displays were presented in which the distractor objects remained in the same locations but white items replaced black items and vice versa. This result suggests that, in this case, observers incorporated object feature information into their mental representations of the repeated displays and, as a result, learning became feature dependent. Finally, Kunar et al. (2006) obtained evidence for contextual cuing in a task in which target location was predicted by the color of the visual display, rather than by the spatial arrangement of objects, although the emergence of this learning effect depended on explicit awareness that color was relevant to search. Thus, according to work with simple stimulus arrays, when color can be profitably used to help recognize contextual information, it becomes an important player in learning.

Although search for targets embedded in simple arrays of shapes and letters may be sensitive to color information, the role that color may play in referencing target positions when real-world scenes serve as the learning context is not clear. On the one hand, there is reason to predict that the color information provided in a scene is used to help reference target locations. Color quality has been shown to affect the ease with which objects and scenes are identified and remembered. For example, memory for color

photographs is superior to memory for black-and-white (Homa & Viera, 1988) or abnormally colored (Wichmann, Sharpe, & Gegenfurtner, 2002) photographs. These effects of color hold for photographs of both natural and artificial objects, suggesting that the effect is independent of an object's or a scene's color diagnosticity—the degree to which color is associated with an object or scene (e.g., bananas are yellow and jungles are green, but sofas and living rooms can be any color). Together, these results suggest that color is used in episodic memory as a cue for recognizing specific exemplars of objects or scenes. On the basis of these findings, one could expect manipulations of color information to influence observers' sensitivity to scene–context covariation.

On the other hand, there are at least two reasons why contextual cuing may rely less on color in real-world scenes than in arrays of letters and shapes. First, as compared with real-world scenes, arrays of shapes and letters are rather impoverished and homogeneous. As such, surface features such as color are highly salient and may, therefore, be very useful when one context is discriminated from another. Scenes, however, are more complex. As a result, many different visual aspects of the scene may contribute to contextual cuing, without one playing a dominant role. Manipulations of color information may, therefore, be overcome by consistencies that remain across other features of the scene. Second, changing the color of a letter may subjectively constitute a change in its identity. As a result, such contexts that incorporate color changes may not be recognized as being the same collection of objects and may, therefore, be treated as novel. Changing the color of an object in a scene may alter its physical appearance but may not change its semantic identity. Thus, if cuing in scenes is influenced primarily by semantic information, changes of color may not have a strong effect on search. Indeed, previous research has shown that the ease with which objects and scenes are recognized affects the speed of learning in contextual-cuing paradigms. As compared with upright scenes, the rate of learning in inverted scenes (which are harder to interpret) is much slower (Brockmole & Henderson, 2006b), and chess experts are able to learn context–target covariation within a chess board much more quickly than novices (Brockmole, Hambrick, Windisch, & Henderson, *in press*), suggesting that semantic identification can play a critical role in the learning of context–target associations.

Across four experiments, we examined the degree to which color information is important in contextual cuing in real-world scenes. Observers searched for letter targets embedded in photographs of natural scenes. These targets were used because their locations could not be predicted a priori from scene identity. Following typical contextual-cuing procedures, amid a series of filler trials on which a scene was presented that had not been shown previously, some scenes were presented repeatedly. In each experiment, we compared the strength of contextual cuing in normally colored scenes with that in scenes in which the color schemes were altered so that they would be considered odd or unnatural (e.g., a neon pink sky). In Experiment 1, we focused on whether the rates of learning

would differ across normal and abnormal color versions of scenes. In Experiment 2, we examined whether learning obtained under one type of color scheme would transfer to different color schemes. In Experiments 3 and 4, we focused on the role of color when color-diagnostic scenes serve as the learning context.

EXPERIMENT 1

In Experiment 1, we compared learning rates when repeated scenes were normally colored and when they were presented in digitally altered color schemes that made them look odd or unnatural (e.g., pink skies). If color is an important feature used in guiding attention in a visual search task, contextual cuing should develop more slowly when the scenes appear in unnatural colors than when they appear in natural colors. In addition, we varied the association between color and target placement across repeated scenes. Repeated scenes were presented either in the same color scheme with each repetition or in a different and unique color scheme each time. If color is an important feature for guiding visual search in the contextual-cuing paradigm, learning rates should be slower in the variable color condition than in the consistent color condition.

Method

Participants. Forty-eight undergraduates from the University of Edinburgh participated for a cash payment. They reported normal or corrected-to-normal vision, with no color vision deficiencies.

Stimuli and Apparatus. The stimuli consisted of 136 full-color photographs of real-world scenes. In addition to the original, 15 unique alternate color versions of each photograph were created. The unedited versions were the same photographs used by Brockmole and Henderson (2006b, Experiment 1; 2006a). The altered-color versions of the scenes were produced through a series of pixelwise manipulations in CIE $L^*a^*b^*$ color space, which represents any color independently of luminance along two color-opponent dimensions: a^* extends from green to red, and b^* from blue to yellow. By manipulating color within $L^*a^*b^*$ color space, we were able to change the colors of the scenes without affecting physical luminance levels (L^*). The color manipulations also preserved the distance within the color space between any given pair of pixels in the image, which means that the difference in color between objects, as well as the variability of color across the image, was unaffected. Seven transformed scenes were produced using the *axis swap* (e.g., changing yellow to red) and *axis invert* (e.g., changing yellow to blue) operations described by Oliva and Schyns (2000). These 7 altered versions and the original version of the scene were then further transformed by rotating the color axes 45° in $L^*a^*b^*$ color space. Figure 1 illustrates an example scene in each of the 16 possible color versions.

Scenes were displayed at a resolution of 800×600 pixels \times 24 bits of color on a 19-in. CRT with a refresh rate of 100 Hz. A single gray T or L subtending 5 pixels horizontally and 7 pixels vertically was digitally inserted into each scene. These letters constituted search targets, and across scenes, each target type occurred equally often and was uniformly distributed across the display. The color, physical luminance, physical contrast, and statistical salience of the targets (Rosenholtz, 2001) were unaffected by the color transformations.¹

Design and Procedure. All the participants viewed a sequence of trials that were divided into 16 blocks. Within each block, eight novel trials and eight repeated trials were randomly intermixed. A novel trial presented a scene that had not been previously shown during the experiment. A repeated trial presented one of eight scenes that had been previously shown. Each of the eight repeated scenes was shown once in each block and was, therefore, repeated 16 times

across the experiment. Within each repeated scene, the target's location was held constant. The target's identity with each repetition, however, was randomly determined.

Participants were equally divided into three between-subjects conditions that differed with regard to the color schemes of the stimuli. In the *consistent color condition*, each scene was presented in 1, and only 1, of its 16 colored versions. For novel trials, the color scheme of each scene was randomly determined. For repeated trials, scenes were presented in one of the 16 possible color versions, and the same version of each repeated scene was used in every block. Across participants, each version of each repeated scene occurred equally often, and no participant saw the same color transformation applied to two different repeated scenes. In the *variable color condition*, novel scenes were again shown in a randomly selected color version, but rather than the same version being used across repetitions, repeated scenes were presented in all of the 16 possible versions, in random order, across the 16 blocks. As a result, the same scene was repeated 16 times, but in a different color scheme each time. For participants in the *control condition*, all the scenes were presented in their original and unedited format. This condition constituted a direct replication of that in Brockmole and Henderson (2006b, Experiment 1).

Trial events were the same in all the conditions. At the beginning of each trial, the participants were shown a blue fixation dot on a gray background and pressed the space bar when ready to view the scene. The participants signaled their detection of the target by pressing one of two buttons on a keyboard, and their reaction time (RT) was recorded. If the participant did not respond within 20 sec, the trial terminated automatically. The participants were given no information about the repetition of the scenes or the color manipulations.

Results

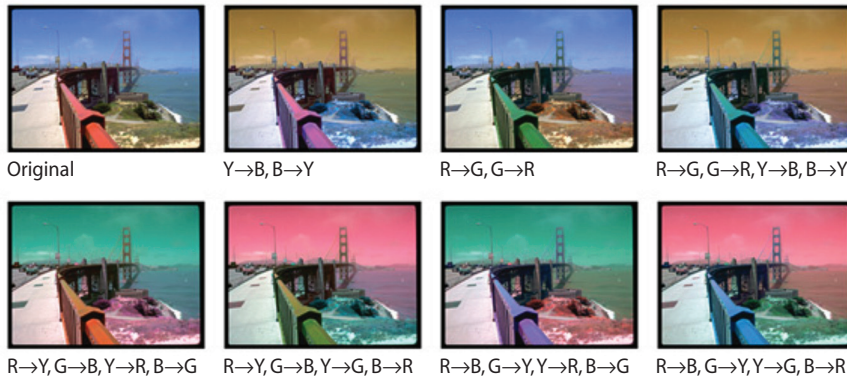
Trials on which RTs were over 20 sec (1% of the trials) or on which an incorrect response was given (2% of the trials) were excluded from further analysis. Analyses also excluded novel trials on which the RT was more than three standard deviations above the participant's mean novel trial RT (1% of the total trials).

Search times are summarized in Table 1, broken down by the experimental factors. An initial analysis focused on search times for the novel trials, since they established the baseline search rate for each condition across epochs. A 3 (condition) \times 16 (block) mixed model ANOVA revealed a reliable effect of block [$F(15,675) = 2.12, p < .01$], but search times did not differ significantly among the three conditions [$F(2,45) = 1.32, p = .28$], nor did these factors interact [$F(30,675) < 1$]. This result indicates that the color transformations did not affect the perceptibility of the targets or task difficulty. Therefore, we calculated the *cuing effect* (Chun & Jiang, 1998) on each block as the difference between novel and repeated trials. These cuing effects are plotted in Figure 2, broken down by condition and block.

A one-way repeated measures ANOVA was conducted within each color condition to demonstrate that the cuing effect increased across blocks, regardless of the color information contained across scenes [consistent color condition, $F(15,225) = 5.06, p < .001$; variable color condition, $F(15,225) = 5.80, p < .001$; control condition, $F(15,225) = 5.19, p < .001$]. Having established that the observers in each color condition were sensitive to the covariation that existed between repeated scenes and the location of targets, a 3 (condition) \times 16 (block) mixed model ANOVA compared the cuing effects across groups. As was expected, a



Color changes from Oliva and Schyns (2000)



+45° rotation

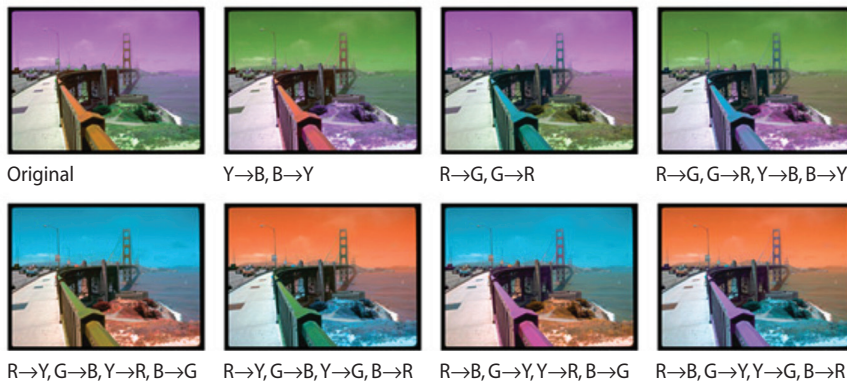


Figure 1. Stimuli used in Experiments 1 and 2. Top: An example scene. The target (a small gray L) is visible just above the bridge's left suspension cable. Bottom: Color transformations were performed in CIE $L^*a^*b^*$ color space, using a combination of axis swap, axis invert, and axis rotation. The captions describe which colors were exchanged in each transformation (R = red, G = green, Y = yellow, B = blue). A full-color illustration is available in the electronic version of this article.

Table 1
Experiment 1 Results: Mean Search Times (in Milliseconds) Broken Down by Color Condition, Block, and Trial Type

Trial Type	Block															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Control Condition																
Novel	3,794	3,922	3,061	2,875	3,143	3,478	3,539	3,321	3,204	3,338	3,068	3,221	3,631	3,600	3,507	2,824
Repeated	3,696	3,222	2,724	2,130	1,871	1,735	1,554	1,468	1,335	1,293	1,363	1,321	1,213	1,251	1,290	1,267
Consistent Color Condition																
Novel	3,884	3,198	3,198	3,426	3,087	3,577	2,854	3,189	2,678	2,996	2,848	3,077	3,259	3,454	2,998	3,208
Repeated	3,281	3,444	2,488	2,044	1,767	1,607	1,518	1,503	1,392	1,384	1,288	1,347	1,323	1,302	1,252	1,166
Variable Color Condition																
Novel	3,888	3,164	3,503	3,404	3,627	3,895	3,327	3,697	3,072	3,265	3,339	3,316	3,365	3,417	3,698	2,996
Repeated	3,855	2,776	2,398	2,209	2,072	1,749	1,591	1,386	1,368	1,380	1,272	1,252	1,214	1,198	1,190	1,163

main effect of block was observed [$F(15,675) = 14.3, p < .001$], since the magnitude of the cuing effect increased over blocks. However, the main effect of color condition was not reliable [$F(2,45) = 1.25, p = .30$], and there was no interaction between color and block [$F(30,675) < 1$]. These results indicate that the differences between novel and repeated trials were unaffected by the nature of the color information available across repetitions of scenes.

Discussion

In all three conditions, search times for repeated scenes decreased over repetitions, which indicates that the observers learned the locations of the targets relative to their contexts. The overall pattern of learning was very similar to that observed by Brockmole and Henderson (2006b). However, there was no difference in learning rates between normally colored scenes, altered-color scenes, and altered-color scenes that changed color on every repetition.

This suggests that the observers did not rely on color information to associate target positions with each scene.

The fact that there was no difference in learning rates between naturally and unnaturally colored scenes is perhaps surprising, given that Wichmann et al. (2002) found that memory for unnaturally colored scenes is worse than memory for naturally colored scenes. However, Wichmann et al. asked participants to distinguish previously viewed photographs from very similar foils drawn from the same semantic category after a 1-sec exposure. In our task, however, the participants' initial search through the scenes (i.e., Block 1) averaged 3.5 sec, and every scene was fairly unique in appearance and meaning. These differences in task and presentation time may underlie the discrepancy pertaining to the importance of color information across the studies. For example, in our experiment, the participants may have been able to make more use of semantic cues in recognizing the repeated scenes, thereby

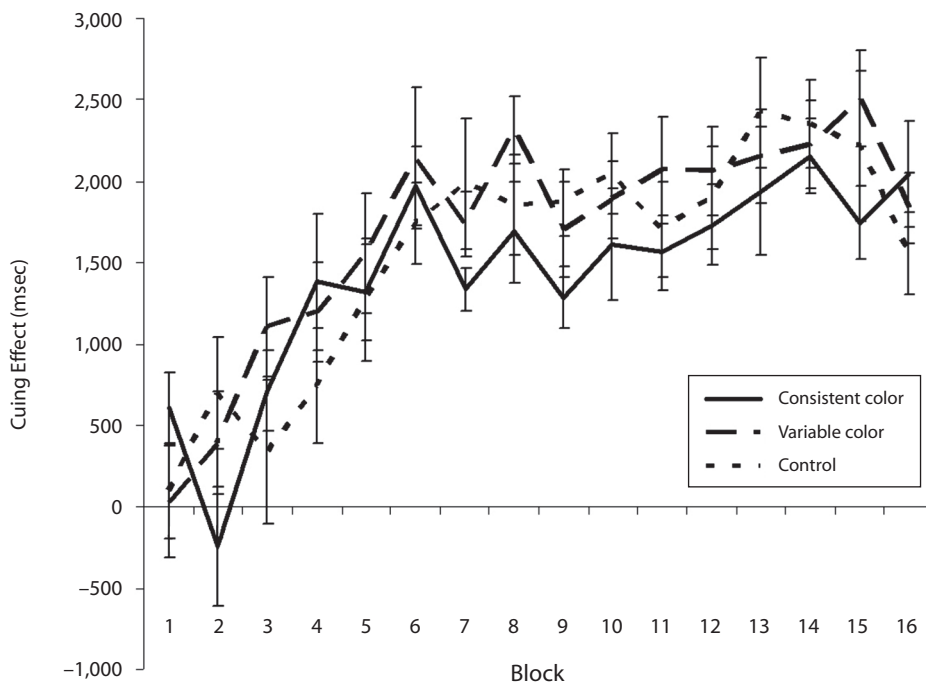


Figure 2. Results of Experiment 1: Mean cuing effects (with standard errors) across blocks, broken down by color condition.

minimizing the importance of color information. We will return to these issues in the General Discussion section.

EXPERIMENT 2

Although the results of Experiment 1 suggest that color is not automatically used during visual search for a consistently located target in a scene, other circumstances exist under which color may be a useful cue. For example, in the variable color condition in Experiment 1, target placement could not be predicted by color information, since the color scheme of repeated trials was changed with each repetition. It is possible, therefore, that people may suppress color information in such cases and, instead, link target position with, for example, the semantic identity of the scenes (see Brockmole et al., 2006; Brockmole et al., in press). If color were, however, to be correlated with target position, observers might use this information to facilitate search, since it is a reliable cue to target location. Such a possibility is suggested by previous research that showed that color was included in contextual representations when it predicted target location but was ignored when it was not predictive (Jiang & Song, 2005).

In Experiment 2, we explored this possibility by examining whether the benefits obtained during learning, under circumstances identical to the consistent unnatural colors condition in Experiment 1, would transfer to situations in which the once consistent color scheme was violated. In Blocks 1–15, each repeated scene was presented in the same color scheme, as in the consistent color condition in Experiment 1. In the 16th and final block, the color scheme of each repeating scene was altered. If predictive colors are included in the contextual representations of scenes, we would expect to see a significant increase in search times (i.e., a reduction in the benefits of contextual cuing) in the final block of the transfer test.

Method

Participants. Sixteen undergraduate volunteers, who had not participated in Experiment 1, were recruited from the University of Edinburgh. The participants reported normal or corrected-to-normal vision, with no color vision deficiencies.

Stimuli and Apparatus. The stimuli and apparatus were the same as those in Experiment 1. For each color version of the repeating scenes, a *color-opposite* version was selected, which was the version obtained by inverting the scene's a^* and b^* axes (equivalent to rotating the scene 180° in $L^*a^*b^*$ color space). This color-opposite version was displayed on the transfer trial. As such, the color change from learning to transfer was maximized, providing the greatest discrepancy in color information possible.

Procedure. The procedure was identical to that in the consistent color condition in Experiment 1 for Blocks 1–15. In Block 16, each repeating scene was replaced with its color-opposite version.

Results

As in Experiment 1, trials on which RTs were over 20 sec (2% of the total trials) or on which an incorrect response was given (3% of the total trials) were excluded from analysis. Analysis also excluded novel trials on which the RT was more than three standard deviations above the participant's mean novel trial RT (1% of the total trials).

Figure 3 shows search times broken down by trial type and block. To establish the presence of a contextual-cuing effect, search times in Blocks 1–15 were first submitted to a 2 (trial type) \times 15 (block) mixed model ANOVA. Consistent with contextual cuing, main effects of trial type [$F(1,15) = 115, p < .001$] and of block [$F(14,210) = 8.12, p < .001$] were observed, in addition to a reliable interaction of these factors [$F(14,210) = 5.90, p < .001$]. Thus, search times within repeated scenes were shorter than those within novel scenes, and this difference was magnified over repetitions. Critically, a paired samples t test revealed no reliable difference in search times between the final learning block (Block 15) and the transfer block (Block 16) [$t(15) < 1$]. Thus, in the transfer block, search times were unaffected by a complete change in the color composition of the learned scenes.

Discussion

Despite the fact that color composition was correlated with target position during learning, when this association was violated in a transfer block, there was no cost in terms of search speed. This result further suggests that color information is not included in the contextual representations that facilitate visual search in real-world scenes, even when color is a reliable cue to target location. It is possible, however, that Experiments 1 and 2 underestimated the importance of color in visual search through real-world scenes. According to some studies, color is useful only in identifying *color-diagnostic scenes*, which are natural scene categories, such as forests or deserts (Oliva & Schyns, 2000). Many of the scenes used in Experiments 1 and 2 were urban and indoor scenes that had no clear color associations. To address this issue, we replicated Experiments 1 and 2 using only color-diagnostic scenes.

EXPERIMENT 3

Experiment 3 was a replication of Experiment 1, using an image set in which all of the scenes came from a single, color-diagnostic category (parks; see Figure 6). Because images from a single semantic category were used, the participants could not rely on cues such as the semantic identity of the scenes or on the particular objects, shapes, or textures in the scenes to guide their search for the targets, since these cues were similar across scenes. By using color-diagnostic images, we hoped to maximize the effects of the color manipulations: If color has any role in real-world scene contextual cuing, it is most likely to be involved in contextual cuing with color-diagnostic scenes.

Three outcomes are possible. First, because identification of color-diagnostic objects and scenes is hindered when color information is absent or altered (Davidoff & Ostergaard, 1988; Oliva & Schyns, 2000; Ostergaard & Davidoff, 1985; Price & Humphreys, 1989; Tanaka & Presnell, 1999; Wurm, Legge, Isenberg, & Luebker, 1993), contextual cuing may be reduced in the color-altered conditions, relative to the normal-colored control condition. Second, despite slower identification, because images are drawn from a single semantic category and

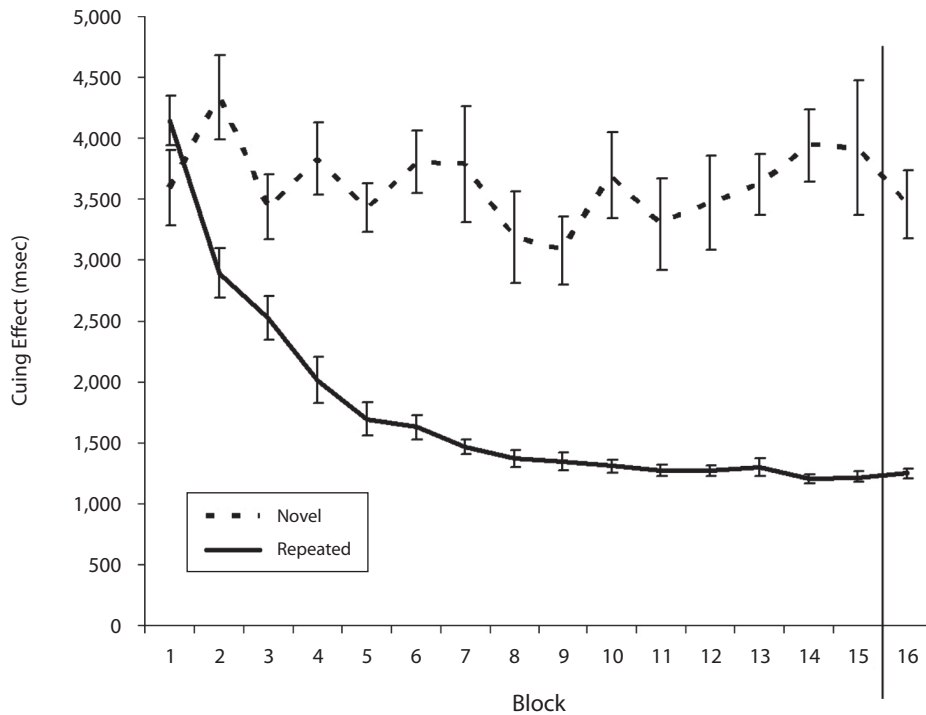


Figure 3. Results of Experiment 2: Mean search times (with standard errors) across blocks for novel and repeated trials. The vertical line separates the learning and the transfer phases of the experiment. Blocks 1–15 are learning blocks, and Block 16 is the transfer block.

have similar color schemes (i.e., primarily green), color alterations to color-diagnostic scenes may increase their perceptual differences. As a result, contextual cuing may be stronger in the consistent color condition than in the variable or normal control conditions, in which color cannot be used to differentiate scenes. Third, contextual cuing could remain impervious to manipulations of color manipulations, since color may not be included in the contextual representations underlying search, even in color-diagnostic scenes.

Assessing Color Diagnosticity

Despite general agreement that color diagnosticity determines whether or not color contributes to object recognition (see Tanaka, Weiskopf, & Williams, 2001, for a review), multiple approaches to operationally defining color diagnosticity have been developed. We used *color consistency* as our measure of color diagnosticity; that is, a scene or object is color diagnostic if it appears in a characteristic color (see Oliva & Schyns, 2000; for other, related approaches, see Tanaka & Presnell, 1999). For example, whereas forests are generally green, deserts are generally yellow, and oceans are generally blue.²

To establish a set of color-diagnostic images for use in the present research, we reasoned that because color-diagnostic scenes have well-defined average colors, the individual scenes in a color-diagnostic category should be very similar in color content to the average for that category. Non-color-diagnostic scenes, on the other hand, should be fairly dissimilar in color to their category aver-

age. Using this criterion, we assessed the color diagnosticity of several categories of real-world scenes. Forests, deserts, and parks were selected as scene categories likely to be color diagnostic. For comparison, we also analyzed city street scenes, which are generally considered non-color-diagnostic, and the image set used in Experiments 1 and 2, which had no unifying category identity and could be expected to be non-color-diagnostic.

The first step in our assessment of diagnosticity was to determine the average color distribution of each image set. For each image, we created color histograms in $L^*a^*b^*$ color space as described in Oliva and Schyns (2000). Average histograms for each image category, shown in Figure 4, were computed by summing the histograms of individual images and then dividing by the number of images in the category. In these histograms, a^* is represented on the horizontal axis and b^* on the vertical axis. The average frequency of each color is represented on a logarithmic scale (see the color bar in Figure 4) in each histogram. Visual inspection of the histograms suggests that the average desert was predominantly yellow and the average forest and park were predominantly green. The average color distributions in the collections of urban and no-category scenes were more normally distributed throughout the color space.

The second step was to calculate a *color histogram distance* for each individual image, as described in the Appendix. The color histogram distance measured the difference between an individual image's color histogram and the average color histogram for that image category.

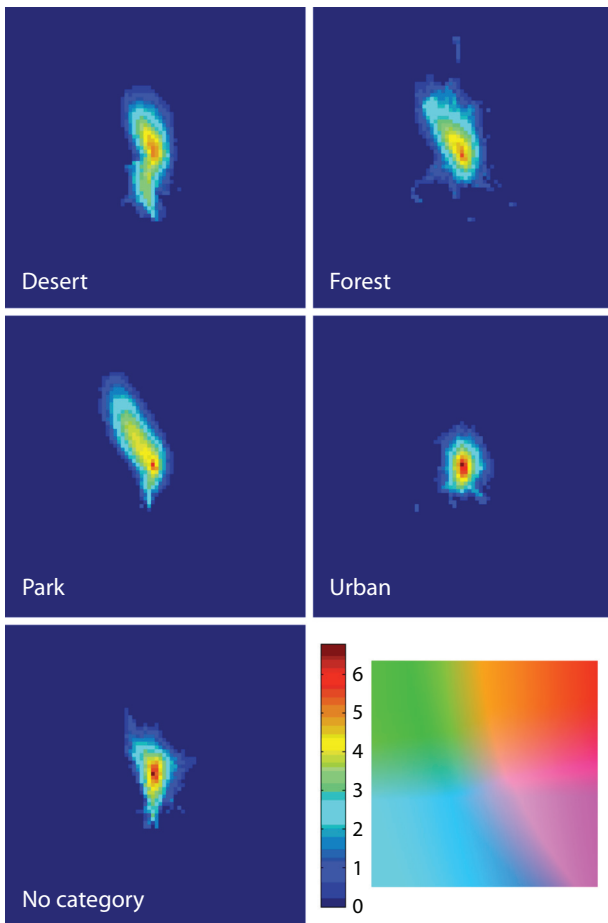


Figure 4. Average color histograms for forests, deserts, parks, urban scenes, and the uncategorized image set used in Experiments 1 and 2, shown on a logarithmic scale. Color histograms are in $L^*a^*b^*$ color space, depicted at the lower right. It is apparent that whereas desert images were predominantly yellow and forests and parks green, the color components in the collections of urban and no-category scenes were more normally distributed. A full-color illustration is available in the electronic version of this article.

Color histogram distances for the scenes in our analysis are shown in Figure 5.

A one-way ANOVA showed reliable effects of scene category [$F(4,675) = 84.5, p < .001$]. Bonferroni-corrected post hoc comparisons showed that each of the three natural scene categories (desert, forest, and park) was significantly different from the urban scenes and the uncategorized scenes in Experiments 1 and 2 ($p < .001$ in all comparisons). The three natural scene categories did not reliably differ from each other, and the urban scenes did not differ significantly from the no-category scenes. From this, we concluded, like others (Oliva & Schyns, 2000), that deserts, forests, and parks are more color diagnostic than urban scenes or the heterogeneous image set used in Experiments 1 and 2. Given their numerically smaller color histogram distance, we chose parks as the color-diagnostic image set to use in Experiments 3 and 4. Example scenes from this image set are illustrated in Figure 6.

Method

Participants. Forty-eight undergraduates, who had not participated in either Experiment 1 or 2, were recruited from the University of Edinburgh and were compensated with a cash payment. The participants reported normal or corrected-to-normal vision, with no color vision deficiencies.

Stimuli and Apparatus. The stimuli were 136 photographs of parks presented at a resolution of 800×600 pixels \times 24 bits of color. A letter target (T or L) was inserted into each image. The targets were placed in unique positions within each image, and across all images, targets were uniformly distributed across the visual field. To select the images for the repeated trials, we conducted a pilot experiment in which 6 observers were shown each scene once and were asked to search for and identify the target. Average search times for each scene were computed, and eight scenes were randomly selected from those falling within the interquartile range of the resulting search time distribution. This procedure ensured that repeated scenes were representative of the entire image set in terms of search difficulty. The apparatus was the same as that in Experiment 1.

Design and Procedure. The design and procedure were the same as those in Experiment 1.

Results

Trials on which RTs were over 20 sec (1% of the total trials), trials on which an incorrect response was given (3% of the total trials), and novel trials more than three standard deviations above the mean (1% of the total trials) were excluded from analysis.

Search times are summarized in Table 2, broken down by the experimental factors. The analyses mirrored those in Experiment 1. An initial 3 (color condition) \times 16 (epoch) mixed model ANOVA on search times through novel trials revealed no reliable effects. Search times did not vary across blocks [$F(15,675) = 1.44, p = .12$] or color conditions [$F(2,45) = 1.10, p = .34$], and the factors did not interact [$F(30,675) < 1$]. Therefore, as in Experiment 1, we calculated the cuing effect on each block as the difference between novel and repeated trials.

Cuing effects across block are illustrated in Figure 7. A one-way repeated measures ANOVA was conducted within each color condition to demonstrate that the cuing effect increased across blocks, regardless of the color information contained across scenes [consistent color condition, $F(15,225) = 9.10, p < .001$; variable color condition, $F(15,225) = 8.37, p < .001$; control condition, $F(15,225) = 7.98, p < .001$]. Having established that the observers in each color condition were sensitive to the covariation that existed between repeated scenes and the location of targets, a 3 (condition) \times 16 (block) mixed model ANOVA compared the cuing effects across groups. As was expected, a main effect of block was observed [$F(15,675) = 24.26, p < .001$], since the magnitude of the cuing effect increased over blocks. However, the main effect of color condition was not reliable [$F(2,45) = 1.81, p = .18$], and there was no interaction between color and block [$F(30,675) < 1$]. As in Experiment 1, even with color-diagnostic scenes serving as the search context, the differences between novel and repeated trials were unaffected by the nature of the color information available across repetitions of scenes.

Discussion

If color is used to facilitate visual search in real-world scenes, it will be most likely to matter in color-diagnostic

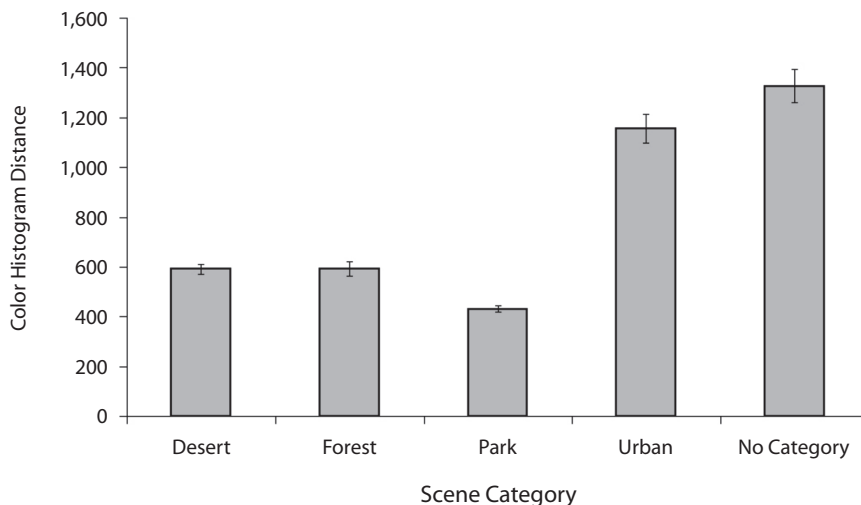


Figure 5. Average color histogram differences (with standard errors) for various image categories.

scenes, because these are the scenes in which color is most likely to play a role in scene recognition. However, the pattern of results with the color-diagnostic scenes matched the results from a heterogeneous image set: Learning rates were the same for naturally colored scenes, consistent unnaturally colored scenes, and variable unnaturally colored scenes.

As in Experiment 1, the observers may have learned in the early blocks of the variable color condition that scene colors were not predictive of target location and may have suppressed the color information. Although one could argue that observers would have had difficulty ignoring the color information in color-diagnostic scenes, to test this possibility, we replicated Experiment 2, the transfer test, using the color-diagnostic parks image set.

EXPERIMENT 4

Experiment 4 was a replication of Experiment 2, using the color-diagnostic parks image set developed for Experiment 3.

Method

Participants. Sixteen undergraduates, who had not participated in any of the previous experiments, were recruited from the University of Edinburgh community and were compensated with a cash payment. The participants reported normal or corrected-to-normal vision, with no color vision deficiencies.

Stimuli and Apparatus. The stimuli and apparatus were the same as in Experiment 3. As in Experiment 2, a *color-opposite* version of each repeated scene was selected for display on the transfer trial.

Design and Procedure. The design and procedure were the same as those in Experiment 2.

Results

Trials on which RTs were over 20 sec (1% of the total trials), trials on which an incorrect response was given (3% of the total trials), and novel trials more than three standard deviations above the mean (1% of the total trials) were excluded from analysis.

Figure 8 shows search times broken down by trial type and block. The analyses mirrored those in Experiment 2. To establish the presence of a contextual-cuing effect, search times in Blocks 1–15 were first submitted to a 2 (trial type) \times 15 (block) mixed measures ANOVA. Consistent with contextual cuing, main effects of trial type [$F(1,15) = 35.6, p < .001$] and of block [$F(15,225) = 9.80, p < .001$] were observed, in addition to a reliable interaction of these factors [$F(15,225) = 8.46, p < .001$]. Thus, search times within repeated scenes were shorter than those within novel scenes, and this difference was magnified over repetitions. Critically, a paired samples *t* test showed no significant difference in search times between the final learning block (Block 15) and the transfer block (Block 16) [$t(15) < 1$]. Thus, in the transfer block, search times were unaffected by a complete change in the color composition of the learned scenes.

Discussion

As in Experiment 2, the complete change of repeated scene colors between learning and transfer blocks had no effect on search times. The participants had learned to associate the targets with the scenes but apparently did not use any stored color information about the scene to facilitate search. This occurred despite the fact that the scenes were color diagnostic and the colors were a reliable predictor of target position.

GENERAL DISCUSSION

The role of color in perception has been debated, with some evidence suggesting that it is used only in low-level vision and other evidence showing that it is important in high-level processes of object identification (see Tanaka et al., 2001, for a review). In the present study, we examined whether contextual cuing within real-world scenes is influenced by the color content of the search context. Observers learned the position of a target letter arbitrarily inserted into



Figure 6. Example (normal color) scenes used in Experiments 3 and 4. Color transformations paralleled those in Experiments 1 and 2 (refer to Figure 1). A full-color illustration is available in the electronic version of this article.

real-world scenes. Amid a sequence of filler trials that presented a scene that had not been shown previously, a subset of these scenes (and target positions) were consistently repeated. In Experiment 1, the observers were shown either unaltered photographs depicting the natural color scheme of each scene or altered photographs in which the color space of the image was manipulated in $L^*a^*b^*$ color space. On repeated trials, either the same color scheme or a different color scheme was used on each presentation. Color consistency, either with respect to the real world or with respect to previous views, did not affect the speed with which scene–target covariation was learned or the magnitude of the search benefit. Experiment 2 then showed that no temporal search costs were generated in a surprise transfer block in which a sudden and significant change to a scene’s color scheme occurred. Experiments 3 and 4 demonstrated that this lack of influence of color on contextual-cuing effects applies to color-diagnostic scenes as well.

The lack of color effects in contextual cuing contrasts with the role of color in other tests of scene memory. For example, previous studies have shown that scene identification and recognition—processes critical to the development of contextual-cuing effects—are impaired when scenes are presented in grayscale or in unnatural colors. These results would seem to suggest that color is used in episodic memory as a cue for recognizing specific exemplars of objects or scenes (Homa & Viera, 1988; Wichmann et al., 2002). So why were we unable to find a role for color information in search for known targets in real-world scenes? Although additional experimentation will be needed to fully resolve this issue, here we outline three possible explanations consistent with the present results and other related findings.

One possibility is that the effects of color on scene identification and recognition are limited to situations in which naming or judgment tasks are used that require explicit re-

Table 2
Experiment 3 Results: Mean Search Times (in Milliseconds) Broken Down by Color Condition, Block, and Trial Type

Trial Type	Block															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Control Condition																
Novel	4,318	4,031	3,814	4,037	3,918	3,775	3,951	3,498	4,145	3,634	4,265	3,377	4,442	3,725	3,779	3,671
Repeated	5,667	4,639	3,618	3,325	2,910	2,567	2,712	2,374	2,354	2,350	2,237	2,169	1,857	1,815	2,094	1,640
Consistent Color Condition																
Novel	4,564	4,130	4,339	4,233	4,751	4,346	4,132	3,965	4,280	4,173	4,496	4,181	4,381	4,551	4,392	3,580
Repeated	5,939	5,148	4,300	4,008	4,219	2,920	2,681	2,161	2,400	2,314	2,025	2,338	2,173	2,049	1,754	2,082
Variable Color Condition																
Novel	4,160	4,259	4,225	4,115	4,361	4,264	4,278	4,155	4,398	3,967	4,692	3,887	4,143	3,855	4,307	3,561
Repeated	6,052	4,058	3,547	3,083	2,988	2,516	2,478	2,306	1,938	1,944	1,955	1,542	1,673	1,537	1,524	1,503

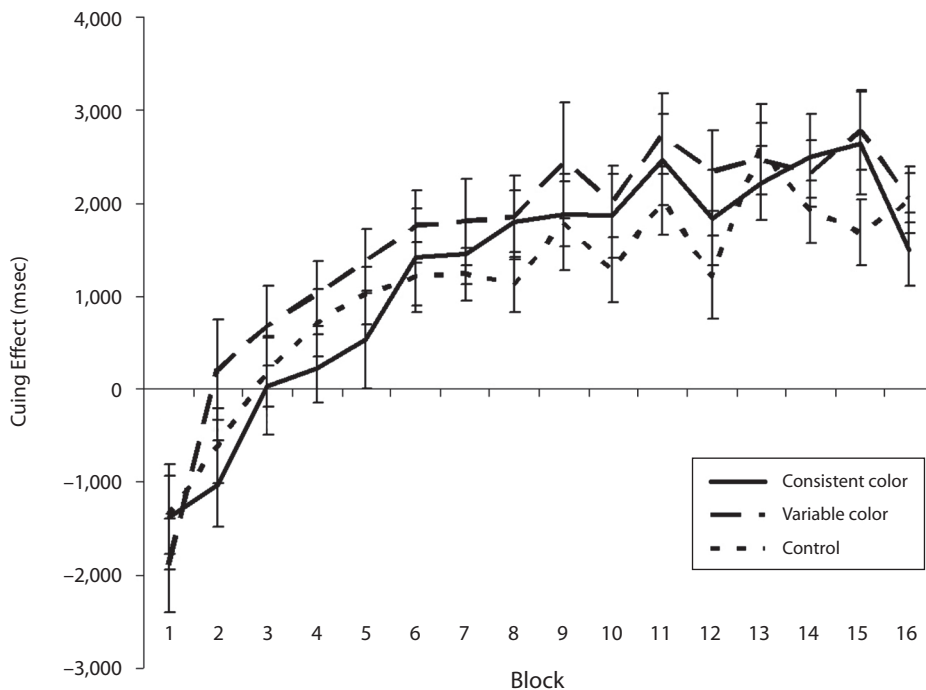


Figure 7. Results of Experiment 3: Mean cuing effects (with standard errors) across blocks, broken down by color condition.

ports from the observer (e.g., Davidoff & Ostergaard, 1988; Ostergaard & Davidoff, 1985; but see Tanaka & Presnell, 1999). In the present study, the observers were not required to explicitly identify the presented scenes. This raises the possibility that explicit reports of scene identity may incorporate specific biases that lead to the conclusion that color plays a role in scene identification and recognition. It may also be the case that the effects of color on scene identification and recognition occur too early in processing to meaningfully influence visually guided behavior in protracted and difficult serial search tasks. Although color can influence the determination of scene gist within 80 msec of scene onset (Castelhano & Henderson, 2008), the present results suggest that this influence may wane over time. Systematic determination of the manner in which color interacts with task demands and the time course over which color can influence visually guided behavior is, therefore, an important avenue for future research.

A second possibility is that during the learning of context–target associations, the target position is referenced with respect to a wide range of visual characteristics of scenes. This redundancy would allow contextual cuing to remain unaffected by variation across a subset of these referencing characteristics. Hence, our manipulations of color information may have been overcome by the consistencies that remained across other features of the scene. Indeed, a context representation that is abstracted from the actual colors of a scene may be more useful in real-world settings than one that is strongly based on scene colors. Long-term color information in the real world may be an unreliable cue for recall of context–target associations, since the absolute color of a scene changes dramatically under differ-

ent conditions: sunlight or shadow, day or night, spring or winter, rain or shine. Given these natural variations in color information available in the scene, visual search for known targets may be guided by other visual aspects of the scene, such as global structure and layout (see Brockmole et al., 2006). Whether color information is encoded into contextual representations if it is the only predictive information in the display (cf. Kunar et al., 2006) remains an interesting question for future research.

A third possibility involves the semantic nature of scenes. Unlike search through arrays of letters or novel shapes, search in scenes is guided, at least in part, by scene semantics (Torralba, Oliva, Castelhano, & Henderson, 2006). In the context of contextual cuing, the same holds true, since initial shifts of attention to a known search target are guided by scene identity information (Brockmole & Henderson, 2006a) and one's ability to extract semantic meaning from a display influences the speed and magnitude of learning (Brockmole et al., in press; Brockmole & Henderson, 2006b). Although changing the color of an object in a scene may alter its physical appearance, it does not necessarily change its semantic identity. Thus, if observers rely on semantic information to guide search, changes to color may not have a strong effect on target localization.

In summary, the present study provides additional evidence that contextual cuing in real-world scenes and that in simple arrays operate differently. Previous research using nonscene stimuli has shown a role for color information in contextual cuing. In contrast, contextual cuing in real-world scenes is not influenced by color information. This is true even when color is a reliable predictor of target location and when color would be expected to

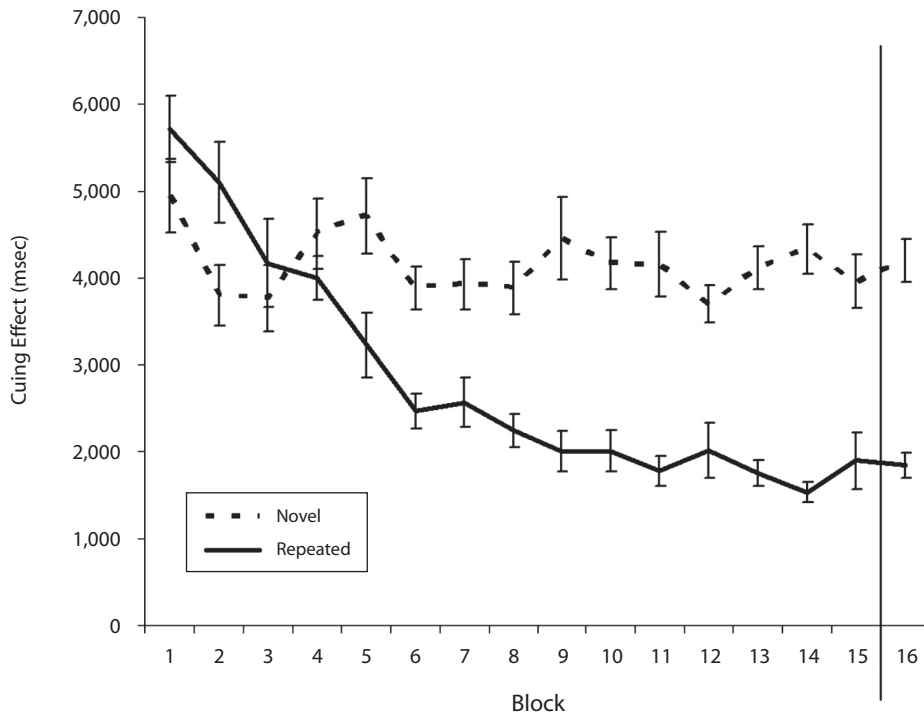


Figure 8. Results of Experiment 4: Mean search times (with standard errors) across blocks for novel and repeated trials. The vertical line separates the learning and the transfer phases of the experiment. Blocks 1–15 are learning blocks, and Block 16 is the transfer block.

aid scene recognition and memory (as is thought to be the case with color-diagnostic scenes). This finding suggests that although color may be particularly important for processing the coarse spatial information that is involved in the fast recognition of scene “gist” (Castelhano & Henderson, 2008; Oliva & Schyns, 2000), it plays less of a default role in guiding search through complex scenes once those scenes have been identified.

AUTHOR NOTE

Portions of this research were conducted by K.A.E. as part of an undergraduate honors project at the University of Edinburgh. We thank Aude Oliva and Ruth Rosenholtz for their discussions pertaining to the manipulation of color, two anonymous reviewers for their helpful comments on a previous version of the manuscript, and Jennifer Cooper, Vivienne McKenzie, and Lynne Grundison for their help with data collection. Address correspondence to J. R. Brockmole, Psychology Department, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland (e-mail: james.brockmole@ed.ac.uk).

Note—Accepted by the editorial board of Editor-Elect Jeremy M. Wolfe.

REFERENCES

- BROCKMOLE, J. R., CASTELHANO, M. S., & HENDERSON, J. M. (2006). Contextual cuing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 699-706.
- BROCKMOLE, J. R., HAMBRICK, D. Z., WINDISCH, D. J., & HENDERSON, J. M. (in press). The role of meaning in contextual cuing: Evidence from chess expertise. *Quarterly Journal of Experimental Psychology*.
- BROCKMOLE, J. R., & HENDERSON, J. M. (2006a). Recognition and attention guidance during contextual cuing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 1177-1187.
- BROCKMOLE, J. R., & HENDERSON, J. M. (2006b). Using real-world scenes as contextual cues for search. *Visual Cognition*, *13*, 99-108.
- CASTELHANO, M. S., & HENDERSON, J. M. (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology: Human Perception & Performance*, *34*, 660-675.
- CHUA, K.-P., & CHUN, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception & Psychophysics*, *65*, 72-80.
- CHUN, M. M. (2000). Contextual cuing of visual attention. *Trends in Cognitive Sciences*, *4*, 170-178.
- CHUN, M. M. (2003). Scene perception and memory. In D. E. Irwin & B. H. Ross (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 42, pp. 79-108). San Diego: Academic Press.
- CHUN, M. M., & JIANG, Y. (1998). Contextual cuing: 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.
- DAVIDOFF, J. B., & OSTERGAARD, A. L. (1988). The role of colour in categorical judgements. *Quarterly Journal of Experimental Psychology*, *40A*, 533-544.
- HOMA, D., & VIERA, C. (1988). Long-term memory for pictures under conditions of thematically related foils. *Memory & Cognition*, *16*, 411-421.
- JIANG, Y. [V.], & CHUN, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology*, *54A*, 1105-1124.
- JIANG, Y. V., MAKOVSKI, T., & SHIM, W. M. (in press). Visual memory for features, conjunctions, objects, and locations. In J. R. Brockmole (Ed.), *The visual world in memory*. Hove, U.K.: Psychology Press.
- JIANG, Y. [V.], & SONG, J.-H. (2005). Hyperspecificity in visual implicit learning: Learning of spatial layout is contingent on item identity. *Journal of Experimental Psychology: Human Perception & Performance*, *31*, 1439-1448.
- JIANG, Y. [V.], & WAGNER, L. C. (2004). What is learned in spatial con-

- textual cuing—configuration or individual locations? *Perception & Psychophysics*, **66**, 454-463.
- KUNAR, M. A., FLUSBERG, S. [J.], HOROWITZ, T. S., & WOLFE, J. M. (2007). Does contextual cuing guide the deployment of attention? *Journal of Experimental Psychology: Human Perception & Performance*, **33**, 816-828.
- KUNAR, M. A., FLUSBERG, S. J., & WOLFE, J. M. (2006). Contextual cuing by global features. *Perception & Psychophysics*, **68**, 1204-1216.
- NIBLACK, C. W., BARBER, R., EQUITZ, W., FLICKNER, M. D., GLASMAN, E. H., PETKOVIC, D., ET AL. (1993). QBIC project: Querying images by content, using color, texture, and shape. *Proceedings of SPIE: Storage & Retrieval for Image & Video Databases*, **1908**, 173-187.
- OLIVA, A., & SCHYNS, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, **41**, 176-210.
- OLSON, I. R., & CHUN, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, **9**, 273-302.
- OSTERGAARD, A. L., & DAVIDOFF, J. B. (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 579-587.
- PETERSON, M. S., & KRAMER, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics*, **63**, 1239-1249.
- PRICE, C. J., & HUMPHREYS, G. W. (1989). The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology*, **41A**, 797-828.
- RÖSENHOLTZ, R. (2001). Search asymmetries? What search asymmetries? *Perception & Psychophysics*, **63**, 476-489.
- TANAKA, J. W., & PRESNELL, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics*, **61**, 1140-1153.
- TANAKA, J. [W.], WEISKOPF, D., & WILLIAMS, P. (2001). The role of color in high-level vision. *Trends in Cognitive Sciences*, **5**, 211-215.
- 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.
- WICHMANN, F. A., SHARPE, L. T., & GEGENFURTNER, K. R. (2002). The contributions of color to recognition memory for natural scenes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 509-520.
- WURM, L. H., LEGGE, G. E., ISENBERG, L. M., & LUEBKER, A. (1993). Color improves object recognition in normal and low vision. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 899-911.

NOTES

1. $L^*a^*b^*$ color space is not a perfect representation for human perceptual color space, so even though physical luminance, contrast, and salience were unaffected by color transformations, human viewers may have experienced some differences. To control for this, the color transformations were counterbalanced (for repeated scenes) or selected at random (for novel scenes). Evidence is provided in the Results section that the color manipulations did not affect the perceptibility of the targets or task difficulty.

2. Oliva and Schyns (2000) employed the additional constraint that color-diagnostic scenes could not overlap in color space. Because we used only one category of scene in these experiments, however, we excluded this criterion here.

APPENDIX

Color histogram distances were computed using an algorithm based on the QBIC definition of color histogram distance (Niblack et al., 1993). For an individual image's $N \times N$ bin color histogram $H(x,y)$ and an average $N \times N$ bin color histogram $A(x,y)$, we first computed a difference histogram, which was simply the difference between H and A :

$$D = |H - A|.$$

We then defined a similarity function, which, for any bin $D(x,y)$, returned the sum of all the bins in the vicinity of $D(x,y)$, weighted by their distance from $D(x,y)$. The weights represented the degree of similarity between $D(x,y)$ and the other bins, ranging from 1 (identical) to 0 (no similarity). We assumed that similarity was a simple function of the distance between two nearby bins in the color histogram. Our similarity function was, therefore, a Gaussian centered on bin $D(x,y)$:

$$\text{Similarity}(x,y) = \sum_{i=1}^N \sum_{j=1}^N D(i,j) * e^{-\frac{(x-i)^2 + (y-j)^2}{2\alpha^2}}.$$

The term α determines the spread of the Gaussian. For the analyses shown in this article, we set α so that the Gaussian would have a value of .5 when the distance between bins was 2. We also ran the analyses using a value of α such that the Gaussian would have a value of .5 when the distance between bins was 4. This did not change our results, with two exceptions: With this broader similarity function, the mean color histogram distance of the no-category scenes was significantly higher than that of the urban scenes ($p < .01$), and the mean color histogram distance of parks was lower than that of deserts ($p < .05$) or forests ($p < .01$).

The color histogram distance between H and A was then computed as

$$\text{dist} = \frac{1}{N^2} \sum_{x=1}^N \sum_{y=1}^N D(x,y) * \text{Similarity}(x,y).$$

In this study, we used 80×80 bin color histograms ($N = 80$).