PSYCHOLOGICAL SCIENCE

Research Report

How Color Enhances Visual Memory for Natural Scenes

Ian Spence, Patrick Wong, Maria Rusan, and Naghmeh Rastegar

University of Toronto, Toronto, Ontario, Canada

ABSTRACT—*We offer a framework for understanding how color operates to improve visual memory for images of the natural environment, and we present an extensive data set that quantifies the contribution of color in the encoding and recognition phases. Using a continuous recognition task with colored and monochrome gray-scale images of natural scenes at short exposure durations, we found that color enhances recognition memory by conferring an advantage during encoding and by strengthening the encoding-specificity effect. Furthermore, because the pattern of performance was similar at all exposure durations, and because form and color are processed in different areas of cortex, the results imply that color must be bound as an integral part of the representation at the earliest stages of processing.*

During the past century, many studies indicated that color plays little or no part in visual memory (for reviews, see Oliva & Schyns, 2000, and Wichmann, Sharpe, & Gegenfurtner, 2002). Most of these experiments used artificial, highly simplified stimuli (line drawings, isolated objects, symbols, or text) that poorly represent the natural world. Recently, Steeves et al. (2004) have emphasized that scene perception can operate independently of object perception, and thus many previous studies may not be directly relevant to the natural role of color. The development of color vision in primates was shaped by adaptation to selection pressures that were part and parcel of a visually complex natural environment. Critical behaviors such as wayfinding, foraging for food, and recognizing predator and prey, or friend and foe, must have played a vital role in the evolution of trichromatic color vision. An evolutionary advantage could have accrued if color had facilitated natural scene recognition by enhancing encoding and recognition.

Recent studies with natural scenes suggest that color is a factor in visual memory (Gegenfurtner & Rieger, 2000; Suzuki & Takahashi, 1997; Wichmann et al., 2002, Experiment 3), although its specific role remains controversial. Each of these studies used a 2 × 2 study-test paradigm: Images of natural scenes were depicted in either color or gray-scale monochrome and were later presented for recognition in either color or monochrome versions. If the results are graphed in the same format (Fig. 1), it becomes clear that there is no common pattern of performance across the four experimental conditions. Although each study found an advantage for the color-color condition, there was considerable variation in the pattern of performance across the other three conditions. As we show later, different patterns imply different roles for color, and thus it is important to establish which of these patterns of variation is accurate. Methodological limitations may have compromised each of the studies, casting doubt on their data and conclusions (see Discussion). Furthermore, because no study has proposed and tested a formal model, how color acts to improve the recognition of natural scenes is still largely an open question.

Color could act to enhance scene recognition in two main ways: (a) by improving edge detection and surface segmentation (Fine, Macleod, & Boynton, 2003) and (b) by being bound as a property of the memorial representation (Clifford, Holcombe, & Pearson, 2004; Rossion & Pourtois, 2004). As an aid to edge detection and surface segmentation, color could be beneficial both at encoding and at recognition. If, in addition, color is bound with form as an integral property of the representation, an encoding-specificity effect should boost performance (Tulving & Thompson, 1973). The encoding-specificity principle asserts that memory is enhanced when the same information available at encoding is also available at retrieval. Encoding operations determine how information is stored, and this, in turn, determines the effectiveness of retrieval cues. Brain areas that are activated during encoding are presumed to be reactivated during retrieval. Because form and color are processed in different areas of cortex, these properties must be synchronized, or bound, if color is to enhance memory. Although the binding of color and form is known to occur relatively rapidly, it is not easy to

---

Address correspondence to Ian Spence, Department of Psychology, University of Toronto, Toronto, Ontario, Canada M5S 3G3, e-mail: spence@psych.utoronto.ca.

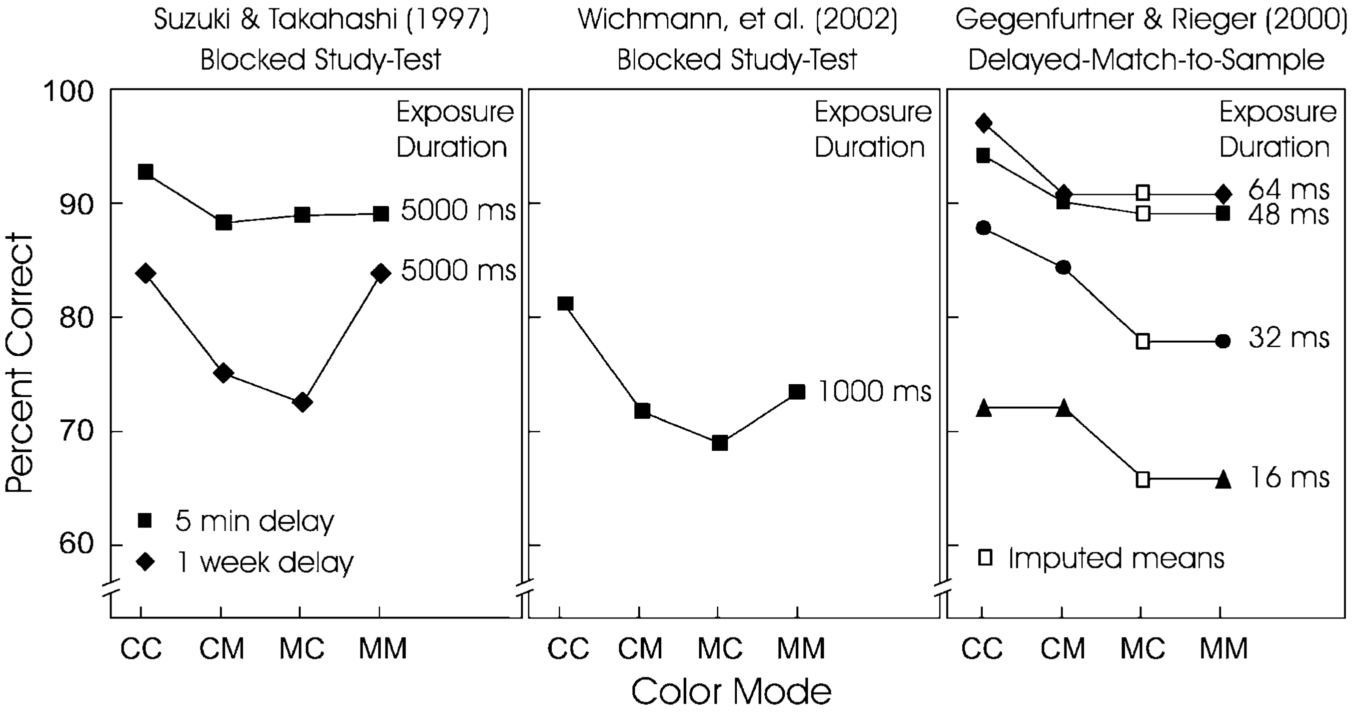

Fig. 1. Mean percentage correct in experiments that used a blocked study-test procedure and a delayed match-to-sample task to investigate the role of color in visual memory. We have redrafted the original graphs for easier comparison with the presentation of our own results. In the color-mode labels, C = color and M = monochrome gray scale; the first letter indicates the nature of the stimuli at encoding, and the second letter indicates the nature of the stimuli at recognition. Gegenfurtner and Rieger (2000) did not report means for condition MC, explaining that performance in this condition ''was generally not different from [performance in] the [MM] condition'' (p. 807). Note that percentages for the three studies are not directly comparable because the tasks were different. It is the shapes of the curves, rather than the absolute success rates, that are revealing.

establish precise lower temporal limits for such binding in all situations (Edwards, Xiao, Keysers, Fő1diák, & Perrett, 2003; Clifford et al., 2004). Failures of binding at short exposures should induce patterns of response different from those obtained at longer exposures for which binding is presumed to be complete.

To clarify the role of color in the 2 × 2 study-test paradigm, we consider seven possible patterns of recognition percentages (Fig. 2) that can be generated by the same linear model that is used in analysis of variance. The model is $\mu_{ij}=\mu + \varepsilon_i + \rho_j + \varepsilon\rho_{ij}$, where μ_{ij} represents mean performance when encoding condition i is combined with retrieval condition j, ε_i represents the contribution of color during encoding condition i, ρ_j represents the contribution of color during retrieval condition j, $\varepsilon\rho_{ij}$ represents the possible nonadditive effect of the idiosyncratic combination ij of encoding and retrieval conditions, and μ is the overall mean level of performance. We have arbitrarily assumed that the sizes of the effects in the model are equal, but deviations from equality do not markedly alter the patterns. Some patterns of performance seem less plausible, a priori, than others. Any pattern that is based on a role for color at recognition, but not at encoding, would seem to be unlikely. Thus, Configurations 2 and 6 in Figure 2 are improbable. Four of the seven patterns (1, 4, 5, and 7) imply a role for color during encoding. Any one of these is plausible given the likely role played by color as an aid to edge detection and segmentation.

EXPERIMENT

Method

We used a continuous recognition procedure that mimics the way in which people see and recognize stimuli in the real world (the rapid serial visual presentation task of Potter, 1976, is similar). Participants—screened to ensure they had normal color vision—viewed (foveally) a sequence of 120 images of natural scenes (Fig. 3) on a monitor. There were 15 participants at each of eight exposure durations ranging from 20 ms to 2,000 ms, for a total of 120 participants. Each scene was presented initially (encoding phase) either in color or in luminance-matched gray-scale monochrome; later, the same scene was presented again either in color or in gray-scale monochrome (recognition phase). A small, centered fixation point (a cross that was 1 pixel wide) was presented for 500 ms before each scene, and a 250-ms mask followed each scene. The mask was identical in size to the stimulus and was composed of randomly rearranged pixels taken from the colored or gray-scale images. Thus, the masks were similar to the stimuli in their luminance and color distributions.

The experimental design was a factorial with two levels of encoding (color vs. monochrome) and two levels of recognition (color vs. monochrome), resulting in four combinations of color modes (color-color, color-monochrome, monochrome-color, and monochrome-monochrome). The encoding and recognition trials

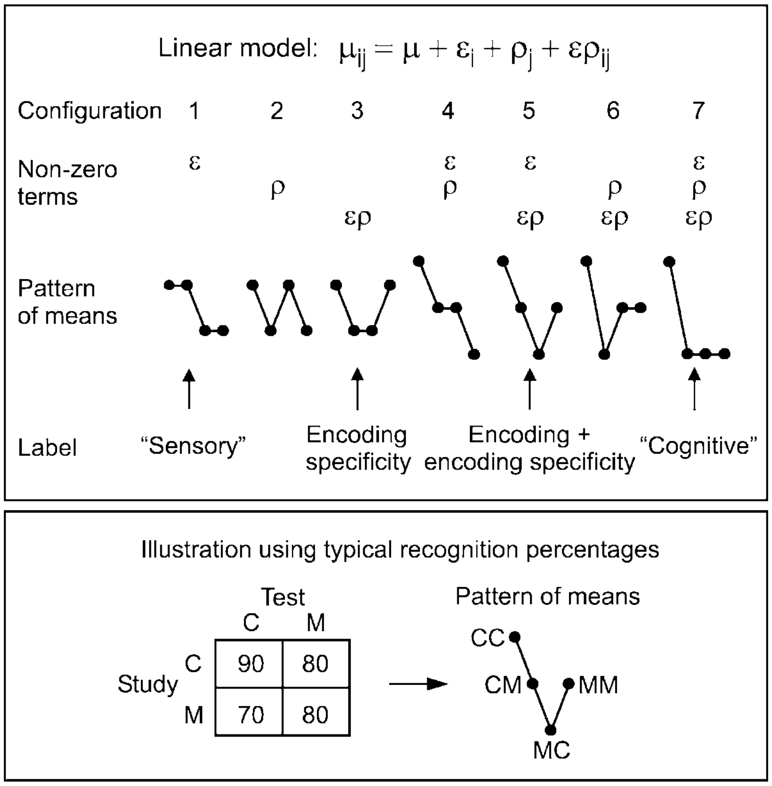

Fig. 2. The seven possible configurations generated by the linear model $\mu_{ij} = \mu + \varepsilon_i + \rho_j + \varepsilon\rho_{ij}$, depending on which model terms have nonzero effect sizes. Each possible pattern of recognition percentages is illustrated with a line graph so that different patterns may be compared easily. The example at the bottom shows a pattern corresponding to particular data obtained in an experiment. C = color and M = monochrome gray scale; where these labels are combined, the first letter indicates the nature of the stimuli at encoding, and the second letter indicates the nature of the stimuli at recognition.

for a given scene were separated by a lag of 1, 2, 4, 8, or 16, and the overall sequence was balanced with respect to each of the five lags and the four color modes. Participants used a four-category rating method, indicating whether each scene was definitely old, probably old, probably new, or definitely new. They made their selection on each recognition trial by pressing one of four keys on a computer keyboard.

Results

Because the data for percentage correct (obtained by collapsing the rating categories) showed the same pattern of results as the d's from signal detection analysis, we discuss percentage correct here (see Fig. 4). The expected effect of exposure duration was very large, $F(7, 112) = 48.6, p < .0001, \omega_{\mathrm{G}}^2 = .382$ (Olejnik & Algina, 2003): Longer exposures were associated with better performance. There was no effect of lag, nor were there interactions of lag with other factors; consequently, the means in Figure 4 are averaged over lags. The standard errors of the means did not vary greatly—they were about 2% at longer exposures and close to 4% at the shorter durations, averaging 3%. The main effect of encoding, $F(1, 112) = 96.3$, $p < .0001$, $\omega_{\mathrm{G}}^2 = .037$, and the interaction of encoding with recognition, $F(1, 112) = 132.5$, $p < .0001$, $\omega_{\mathrm{G}}^2 = .041$, were the only significant sources of variation, with mean squares of comparable size (17,800 vs. 19,700). The mean square for the nonsignificant main effect of color at recognition was comparable in size to the mean square for error variation (220 vs. 140). A comparison between the color-color and monochrome-monochrome conditions showed that color enhanced recognition by about 5% on average, $F(1, 112) = 47.0$, $p < .0001$, $\omega_{\mathrm{G}}^2 = .015$. The same comparison was significant at the .05 level or better for each exposure duration.

Analysis of variance of the response times showed that no speed-accuracy trade-offs were present.

DISCUSSION

Color improves the recognition of natural scenes by about 5%. This is a substantial enhancement. The pattern of performance obtained in this experiment is consistent with Configuration 5 in Figure 2, which implies that color plays a significant role during encoding and also during the recognition matching process. Moreover, the highly similar patterns over the eight different

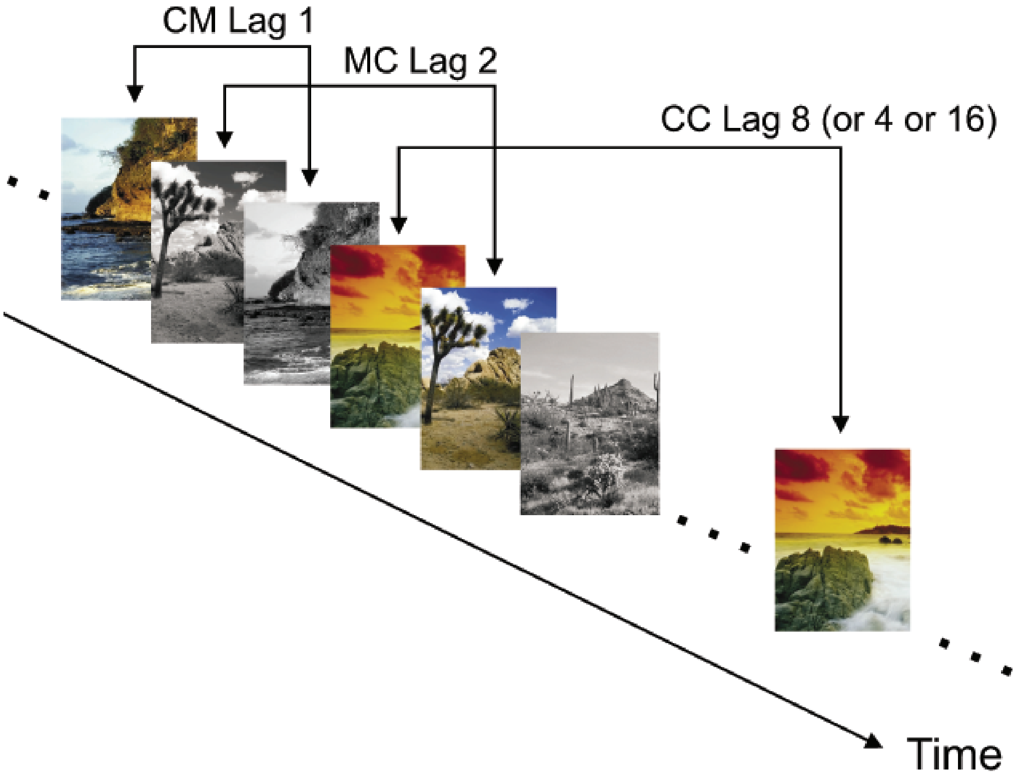

Fig. 3. Excerpt from the trial sequence in the continuous recognition task. The order of the presentation sequence is indicated by the arrow labeled ''Time.'' For clarity, the fixation points and masks presented before and after the images are not shown. C = color and M = monochrome gray scale; the first letter indicates the nature of the stimuli at encoding, and the second letter indicates the nature of the stimuli at recognition. The two-headed arrows call attention to the different lags in the sequence.

exposure durations confirm that encoding and encoding-specificity processes operate in comparable fashion at very short and longer exposure durations. No other study has demonstrated a similar invariance over such a wide range of exposure durations. This result suggests that chromatic information is bound in a visual representation of a natural scene at the very earliest stages of processing—earlier than had previously been supposed.

Performance in the color-color condition was superior, reflecting the role of color during encoding and also during the cued recognition process. Performance in the monochrome-monochrome condition was either equal or superior to that in the color-monochrome and monochrome-color conditions. Although color was not available to enhance edge detection and segmentation during encoding in the monochrome-monochrome condition, the availability of exactly the same form and luminance information at encoding and recognition was more important than the possible augmentation of segmentation processes that color might provide. Although color may have assisted form perception in the color-monochrome condition, color and form are bound at encoding, and, as a result, the lack of color during the recognition phase would have reduced the similarity between the initially encoded representation and the representation constructed at recognition, making a match more difficult. In the monochrome-color condition, color was potentially able to assist with edge detection and segmentation during the recognition phase; however, because this information was not available during encoding, the presence of color at recognition merely served to interfere with the matching process (performance was about 6% worse in the monochrome-color condition than in the monochrome-monochrome condition). Our modeling confirmed that there was no main effect of color at recognition, only an enhanced encoding-specificity effect, which depended on color having been available at encoding.

One study that used separate blocked encoding and recognition phases (Suzuki & Takahashi, 1997) used nonnatural images that contained man-made objects, and this may explain why their results differ from ours. Steeves et al. (2004) have shown that color plays a different role with natural as opposed to nonnatural scenes, and Wichmann et al. (2002, Experiment 1, Fig. 3) found that the performance difference between colored and noncolored scenes is smaller with man-made than with natural scenes. Nonnatural scenes generally exhibit greater variability of color in more homogeneous patches compared with natural scenes, and man-made objects contain sharper, longer, and more regular edges than natural objects; consequently, color is less likely to be useful as an aid to edge detection in non-

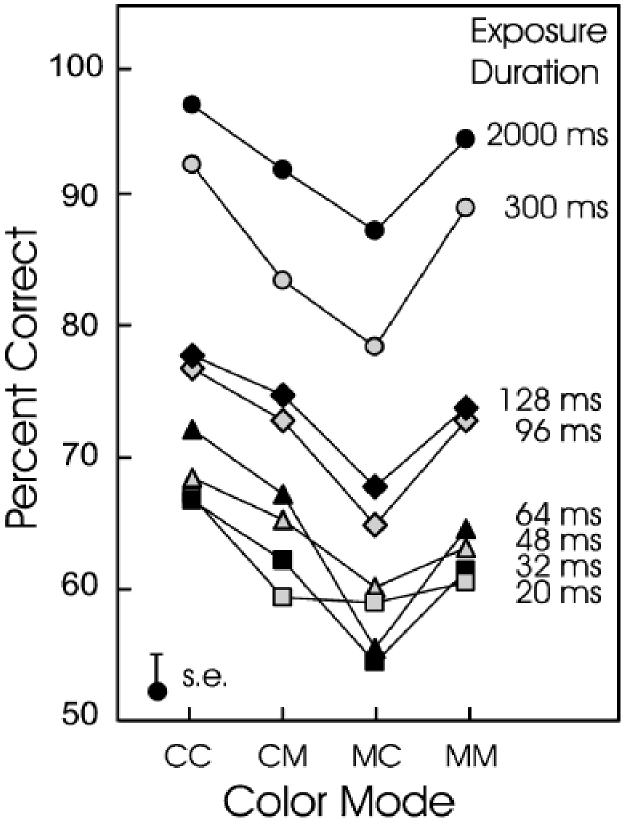

Fig. 4. Mean percentage correct in the continuous recognition task. The average standard error is illustrated by a single symbol to avoid clutter. In the color-mode labels, C = color and M = monochrome gray scale; the first letter indicates the nature of the stimuli at encoding, and the second letter indicates the nature of the stimuli at recognition.

natural scenes. This suggests that a pattern of means like that of Configuration 3 in Figure 2 should be a better fit than Configuration 5 for performance with images of nonnatural scenes; indeed, one pattern that Suzuki and Takahashi (1997) obtained fits Configuration 3 quite well (see our Fig. 1, left panel, lower curve). The pattern of means in another study that used blocked encoding and recognition phases (Wichmann et al., 2002, Experiment 3) is similar to our Configuration 5. Although Wichmann et al. used some images containing nonnatural objects, the majority (75%) of their scenes contained no nonnatural objects.

Gegenfurtner and Rieger (2000) proposed a qualitative shift in how the brain stores images that include color, suggesting (p. 805) that "sensory" processes (Configuration 1) dominate at short durations, whereas "cognitive" processes (Configuration 7) become more involved at longer durations. However, the patterns of means they obtained (see our Fig. 1, right panel) may have been partly the result of an artifact. Participants in their delayed match-to-sample task may have used a focusing strategy by concentrating on a small area of the study image, so that the choice between target and distractor reduced to a decision about that small area. More likely, however, is the analogous possibility that distinctive (bright or strongly colored) areas in the study images captured attention at encoding. In this case, the task would have reduced to the recognition of such repeated anomalies. This artifact would be particularly relevant at short exposure durations, at which it is undoubtedly more difficult to process the whole image. The result would be a spurious change in the pattern of means with increasing exposure times. Our continuous recognition task was resistant to a focusing bias because there were 1, 2, 4, 8, or 16 intervening images between study and test images. Also, our images were presented for exactly the same brief durations at study and test. In Gegenfurtner and Rieger's study, the target and distractor images were available for as long as the participant needed. Finally, some of their images were of nonnatural scenes, whereas all of our images portrayed natural environments.

We found no effect of lag, confirming that visual recognition memory does not degrade rapidly in the short term. In our Lag 1 condition, the delay until the first repeated image was at least 3,000 to 4,000 ms, depending on the participant's previous two decision times. Lags of 2, 4, 8, and 16 created longer delays, with the longest lag producing an interval of more than 30 s. Hence, recognition performance was unchanging between about 4 s and 30 s, which is a longer interval than the period of brief persistence that Potter, Staub, Rado, and O'Connor (2002) have suggested is necessary to form a coherent representation by combining information from several successive fixations.

Because color vision evolved in a natural environment, it is important to study how color assists visual memory in a natural context. Steeves et al. (2004) have discussed how scene perception can operate independently of object perception and have offered neuroimaging evidence for a cortical area that is specialized for viewing scenes. Furthermore, this region is anatomically distinct from areas activated by viewing isolated objects. Our modeling has helped clarify how color plays its role in enhancing memory for natural scenes by creating explicit predictions that could be supported or rejected by empirical testing. Because the sizes of the significant encoding and encoding-specificity effects we obtained were similar, we conclude that these processes are equally important in visual memory for scenes. Although color may facilitate edge detection or segmentation, this appears to be advantageous only during encoding. During recognition, encoding-specificity processes dominate; that is, it is not the presence of color that is important, but rather the quality of the match between the attributes of the initially presented image and the to-be-recognized partner image. Furthermore, the similarity of performance at all exposure durations suggests that color is bound as an integral part of the representation at the earliest stages of processing.

Acknowledgments—This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to the first author.

REFERENCES

Clifford, C.W.G., Holcombe, A.O., & Pearson, J. (2004). Rapid global form binding with loss of associated colors. *Journal of Vision*, *4*, 1090–1101.

Edwards, R., Xiao, D., Keysers, C., Főldiák, P., & Perrett, D. (2003). Color sensitivity of cells responsive to complex stimuli in the temporal cortex. *Journal of Neurophysiology*, *90*, 1245–1256.

Fine, I., Macleod, D.I.A., & Boynton, G.M. (2003). Surface segmentation based on the luminance and color statistics of natural scenes. *Journal of the Optical Society of America*, *20*, 1283–1291.

Gegenfurtner, K.R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, *10*, 805–808.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods*, *8*, 434–447.

Oliva, A., & Schyns, P.G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, *41*, 176–210.

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

Potter, M.C., Staub, A., Rado, J., & O'Connor, D.H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1163–1175.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*, 217–236.

Steeves, J.K.E., Humphrey, G.K., Culham, J.C., Menon, R.S., Milner, A.D., & Goodale, M.A. (2004). Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *Journal of Cognitive Neuroscience*, *16*, 955–965.

Suzuki, K., & Takahashi, R. (1997). Effectiveness of color in picture recognition memory. *Japanese Psychological Research*, *39*, 25–32.

Tulving, E., & Thompson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.

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, and Cognition*, *28*, 509–520.

(Received 2/24/05; Revision accepted 4/26/05;
Final materials received 5/2/05)