

# Color and shape interactions in the recognition of natural scenes by human and monkey observers

**Stefanie Liebe**

Max Planck Institute for Biological Cybernetics,  
Department of Physiology of Cognitive Processes,  
Tübingen, Germany



**Elvira Fischer**

Max Planck Institute for Biological Cybernetics,  
Department of Cognitive and Computational Psychophysics,  
Tübingen, Germany



**Nikos K. Logothetis**

Max Planck Institute for Biological Cybernetics,  
Department of Physiology of Cognitive Processes,  
Tübingen, Germany



**Gregor Rainer**

Max Planck Institute for Biological Cybernetics,  
Department of Physiology of Cognitive Processes,  
Tübingen, Germany, &  
Department of Medicine, University of Fribourg,  
Chemin du Musee 5, Switzerland



Trichromatic color vision is a fundamental aspect of the visual system shared by humans and non-human primates. In human observers, color has been shown to facilitate object identification. However, little is known about the role that color plays in higher level vision of non-human primates. Here, we addressed this question and studied the interaction between luminance- and color-based structural information for the recognition of natural scenes. We present psychophysical data showing that both monkey and human observers equally profited from color when recognizing natural scenes, and they were equally impaired when scenes were manipulated using colored noise. This effect was most prominent for degraded image conditions. By using a specific procedure for stimulus degradation, we found that the improvement as well as the impairment in visual memory performance is due to contribution of image color independent of luminance-based object information. Our results demonstrate that humans as well as non-human primates exploit their sensory ability of color vision to achieve higher performance in visual recognition tasks especially when shape features are degraded.

Keywords: primate vision, color vision, recognition memory, natural scenes

Citation: Liebe, S., Fischer, E., Logothetis, N. K., & Rainer, G. (2009). Color and shape interactions in the recognition of natural scenes by human and monkey observers. *Journal of Vision*, 9(5):14, 1–16, <http://journalofvision.org/9/5/14/>, doi:10.1167/9.5.14.

## Introduction

Surfaces and objects absorb and reflect light of different wavelengths. These variations in spectral composition represent an important source of information in our natural visual environment. Color vision allows the primate visual system to access this type of information. Based exclusively on differences in the spectral distribution, most primates are able to perceive colors of objects and surfaces and could potentially use this information to recognize them.

In the past, it has been suggested that color could improve object and scene recognition in two major ways.

One way by which color could enhance object identification is that color makes the retrieval of information

from memory more efficient. This could be achieved by two different mechanisms. First, color could help the retrieval process by being a property of the mnemonic representation of objects and scenes in long-term memory. Such “color knowledge” leads to improved object recognition and classification (Clifford, Holcombe, & Pearson, 2004; Hanna & Remington, 1996; Naor-Raz, Tarr, & Kersten, 2003; Tanaka, Weiskopf, & Williams, 2001). Supporting evidence comes from studies mostly using single objects and scenes that are associated with a particular color like a yellow banana or a green forest (Hansen, Olkkonen, Walter, & Gegenfurtner, 2006; Oliva & Schyns, 2000; Ostergaard & Davidoff, 1985; Tanaka & Presnell, 1999). For example, Tanaka and Presnell (1999) showed that recognition of objects with high color diagnosticity (i.e., a taxi or a fire engine) benefited more

from color than the recognition of objects with low color diagnosticity (i.e., a table or a dog). Second, color could make the retrieval of information more efficient if there is an overlap of color information at the encoding and the retrieval phase. Evidence in favor of this argument stems mainly from delayed matching-to-sample studies that independently varied the presentation and query color of natural images (Gegenfurtner & Rieger, 2000; Spence, Wong, Rusan, & Rastegar, 2006; Wichmann, Sharpe, & Gegenfurtner, 2002). If the color properties of the firstly presented (sample) stimulus matched those in the stimulus that needs to be matched (test), higher recognition performance was observed. For example, in a delayed matching-to-sample paradigm, Spence et al. (2006) found that recognition performance was higher in a condition in which both the sample and the test stimuli were presented in color, when compared with a condition in which either the sample or test stimuli were colored. The authors concluded that image color promotes recognition since color information was congruent between the sample and the test stimulus in the one but not in the other conditions.

Another way by which color could lead to improved object identification is by augmenting surface segmentation and contour detection at early stages of sensory processing (Fine, MacLeod, & Boynton, 2003; Gegenfurtner & Rieger, 2000; Mullen, Beaudot, & McIlhagga, 2000; Price & Humphreys, 1989; Spence et al., 2006; Wichmann, Braun, & Gegenfurtner, 2006; Wurm, Legge, Isenberg, & Luebker, 1993; Yip & Sinha, 2002). These studies suggest that color helps recognition during the perceptual encoding of a stimulus, irrespective of whether the object that is perceived is associated with a particular color or not. Some of these studies found that color is especially helpful for recognition when shape cues were less informative about an objects' identity (Price & Humphreys, 1989; Yip & Sinha, 2002). These findings imply that the interaction between luminance- and color-based structural information is critical in answering the question whether color is important for object recognition. However, the question to what extent color interacts with luminance-based shape segmentation in order to improve object recognition cannot be readily answered: None of the studies examined the relationship between shape and color degradation in a systematic fashion.

In our study, we investigate two different aspects regarding the role of color for recognition memory. We employ a stimulus degradation procedure based on Fourier decomposition that allows us to assess the separate contribution of color on recognition aside from shape information. The hypothesis that color helps surface segmentation implies a similar advantage of color on recognition in other primate trichromatic species that do not necessarily possess a similar semantic knowledge system as that of the human. To get insights on how color affects recognition memory in the human compared to the non-human primate, we studied the effect of color on recognition in rhesus monkeys. A large body of evidence

suggests that the early stages of sensory and perceptual processing underlying color vision are very similar in both species (Croner & Albright, 1999; Gegenfurtner & Kiper, 2003; Jacobs, 1996; Koida & Komatsu, 2007; Kusunoki, Moutoussis, & Zeki, 2006). However, whether color has an effect on recognition in this species has not been addressed yet.

Figure 1 illustrates the different color conditions for a sample stimulus that was used in the recognition experiments carried out with human and monkey observers. By linear interpolation with a random phase mask (i.e., varying the phase coherence between an original image and a noise pattern), we were able to parametrically alter the amount of shape and color information that was contained in natural images. The manipulation of the phase of natural images has been shown to be an effective tool for studying perception and recognition of natural images (Oppenheim & Lim, 1981; Piotrowski & Campbell, 1982; Wichmann et al., 2006). Our method has also been used in several previous studies (Rainer, Lee, & Logothetis, 2004; Rainer & Miller, 2000). In this study, we ensured that the rms contrast remained constant throughout the stimulus conditions (Dakin, Hess, Ledgeway, & Achtman, 2002; see also [Supplementary Materials](#)).

In condition a (“achromatic condition”), we mixed achromatic noise with achromatic images. Here, shape information in the stimuli was solely based on luminance cues. In condition b (“image-specific color condition”), we mixed achromatic noise with colored natural images. Whereas in condition a only luminance information can be the basis for recognition, in condition b natural image color can also contribute to this process. At a given coherence level, luminance-based shape information is equal for conditions a, b, and c. Thus, a change in performance between these two conditions can only be attributed to a contribution of color to recognition performance. In a third condition c (“colored noise condition”), we mixed chromatic noise and achromatic natural images. Here, color is present in form of visual noise. Since at a given coherence level luminance-based shape information is equal for conditions a and c, this control condition allows us to assess whether color that does not carry information about objects might be able to impair recognition performance by interfering with luminance-based shape recognition.

Finally, it is important to note that color is a very salient visual feature (Davidoff, 1991; Wichmann et al., 2002). Color can therefore draw the attention of the observer to informative or non-informative regions in an image and thus promote or interfere with recognition performance. To assess possible attention-related effects of color for recognition, we employed stimulus condition d (“color-color condition”) in which a colored noise mask is interpolated with a colored image. Here, color can be informative (image color) as well as non-informative (noise color) at the same time. Thus, any positive or negative effects of color on attention/recognition should

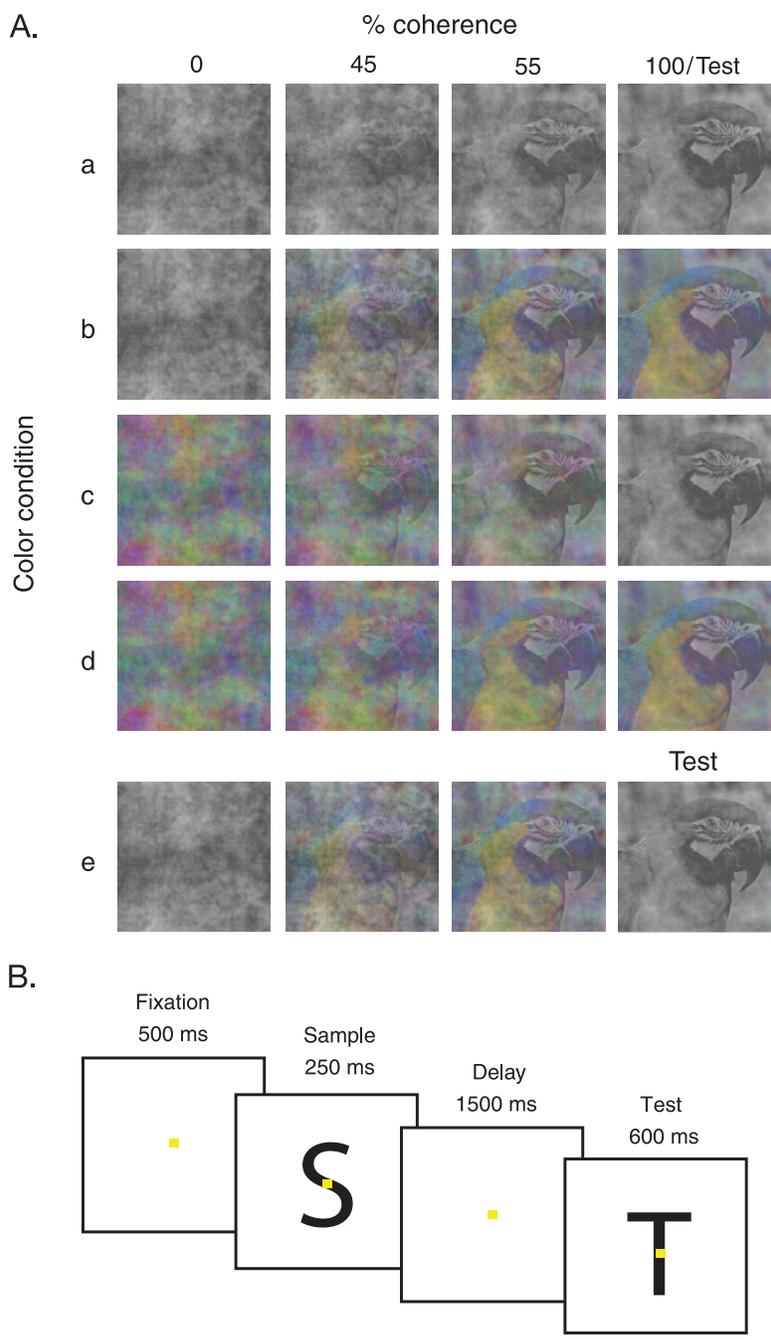


Figure 1. Behavioral task and stimulus conditions for a sample image. (A) Sample stimulus shown at various stimulus conditions used in the present experiment. Each row depicts a combination between a noise pattern (0% coherence) and a supra-threshold natural image (100% coherence) at intermediate levels (45% and 55% coherence); (a) achromatic condition, an achromatic noise mask is interpolated with an achromatic natural image; (b) image-specific color condition, an achromatic noise mask is interpolated with a colored natural image; (c) colored noise condition, a chromatic noise mask is interpolated with an achromatic natural image; (d) “color–color” condition, a colored noise mask is interpolated with a colored stimulus; (e) image-specific color condition with achromatic test stimulus; here an achromatic noise mask is interpolated with a colored image for the sample, the test is always the respective achromatic version. (B) The sequence of trial events and respective time of events. After fixation, a sample stimulus is presented for 250 ms. This is followed by a delay period of 1500 ms during which observers hold fixation and no stimulus is shown. After the delay period, a test stimulus is shown. Human subjects were instructed to press a button whenever the test stimulus appeared to match the sample. Monkeys were rewarded when releasing the lever to match trials and withholding the response in non-match trials; 50% of the trials are match trials, 50% are non-match trials.

be cancel out. By comparing recognition performance in this condition to the achromatic condition, we can test for specific, i.e., attention-independent effects of color for recognition.

Lastly, it has been argued that color can promote the retrieval of information from memory. To test this possibility, we introduced stimulus condition e. Here, we interpolate an achromatic noise mask with a colored image like in condition b. However, in condition e only the sample stimulus is colored whereas the test stimulus is always shown in the respective achromatic version. Thus, color can play a role during the encoding of the scene but not during retrieval. A comparison of condition e with condition b will therefore investigate possible retrieval-related effects of color for recognition memory.

## Experiments with human observers

### Methods

Eleven volunteers (6 males, 5 females, age 23–32) participated in an experiment testing stimulus conditions a, b, and c. We also conducted additional experiments involving conditions d and e with four human observers (3 males, 1 female, age 25–32). All subjects were paid for participation in this study, which was approved by the local ethics committee (MPI). They were all in good health and had normal or corrected-to-normal visual acuity and normal color vision.

Subjects were seated in front of a computer screen at a distance of approximately 58 cm. Stimuli were  $7^\circ \times 7^\circ$  in size, with 24-bit color depth and presented at the center of gaze on a monitor (Intergraph 21sd107) with linear luminance response (gamma corrected). Subjects were instructed about the experimental procedures and underwent several training trials prior to the start of the experiment. The behavioral task of the human observers was a delayed matching to sample (DMS). On each trial, the subjects initially had to fixate a small fixation cross at the center of the screen. After a 200-ms fixation period, a sample image was presented for 250 ms at any of the possible stimulus conditions. After a 1500-ms delay period, the test stimulus was presented for 600 ms. Subjects were instructed to press a button whenever the test image matched the sample stimulus or to withhold the button press in the opposite case. In the case of the 0% coherence level (i.e., being equal to a noise pattern only), the subjects were asked to guess. In one experimental session, all stimulus conditions were randomized, match and non-match trials appeared equally often within a total of 360 trials. A set of 20 natural images was chosen for the human experiments. All images were novel to the subjects. The stimuli that were used in all of the

experiments were chosen from the Corel-Photo-CD “Corel Professional Photos” comprising a collection of natural images showing birds, flowers, monkeys, and butterflies in their natural surroundings. The images used in this study were randomly selected. All images were manipulated by Fourier techniques that have been described in detail elsewhere (Rainer et al., 2004). In short, the images were first normalized to have identical amplitude spectra in Fourier space resulting in matched spatial frequency content and contrast. We parametrically varied the amount of visual noise by combining the Fourier phase spectra of the natural images with a random phase spectrum using the inverse Fourier transform, at six coherence levels (0%, 25%, 35%, 45%, 55%, and 100%) to obtain the pure noise, intermediate noise, and full image conditions. This procedure was applied to each RGB channel independently. To create achromatic images, we converted the colored natural images as well as the colored noise masks to luminance-matched grayscale images using the weighted sum of the different RGB components using weights that were chosen based on luminance measurements of the monitor (weighting factors for R: 0.2458, G: 0.6475, and B: 0.1069). All images were normalized to mean intensity of 0.5 and rms contrast of 0.033 in a range of [0;1]. The space-averaged mean luminance was approximately  $37.7 \text{ cd/m}^2$ . For each session, we used different random phase masks to generate the images. All images shown were matched in overall luminance contrast but could potentially vary locally in this measure. We quantified local luminance contrast for each degradation level and each color condition. [Supplementary Figure 3](#) shows the distribution of differences in local rms luminance contrast from image patches with a size of  $70 \times 70$  pixels, which corresponded approximately to an area of  $2.6 \times 2.6$  degrees of visual angle. At any given coherence level, no significant differences in local root-mean-square contrast were found between identical images patches in each of the three different color conditions used. Thus, at a given coherence level, all images contained equivalent form or shape information based on (local) luminance cues and only differed in their respective color content. To compare the distribution and magnitude of color content across the two color conditions, we transformed the images into the cone-opponent color (LMS) space, in which one axis describes the difference in the activation between the L- and M-cones ( $L - M$ ), and the other axis describes the difference between the S- and the sum of the L and M cones ( $2S - (L + M)$ ). We calculated the LMS responses based on emission spectra of the monitor phosphors as measured by the Minolta CRT Color Analyzer CA-100 and the absorption spectra of LMS cones given by Smith and Pokorny (1975). [Supplementary Figure 4A](#) shows the distribution of chromaticity samples expressed as % cone contrast from image patch samples of a sample image at various color conditions. Chromaticities are equally spread across color space and have similar magnitudes between the two color conditions (condition

b, colored image vs. condition c, colored noise condition) when comparing variance and mean of chromaticity distributions at the following coherence levels: 100% colored image (full image) vs. 0% colored noise (pure colored noise mask), 55% colored image vs. 45% colored noise, 45% colored image vs. 55% colored noise. [Supplementary Figure 2B](#) illustrates that images were also matched when comparing the magnitude (2B) and distribution of average chromaticity across all images used.

In order to derive the psychometric function, we firstly calculated psychophysical performance (% correct responses) at each coherence level for each subject individually. We subsequently averaged recognition performance across the subjects and fitted a psychometric function (logistic function fit) using the `psignifit` toolbox version 2.5.6 for Matlab, which implements the maximum likelihood method described by Wichmann and Hill (2001a, 2001b).

In order to do statistical comparisons between threshold estimates of different conditions, we obtained estimates for parameters of interest for psychophysical functions (slopes, thresholds) through a bootstrapping procedure using the bootstrap method with BCA correction implemented in `psignifit`, based on 2000 simulations. This method is also described in detail in Wichmann and Hill (2001b). This procedure provides 95% confidence intervals for the parameter estimates of interest. We reported threshold estimates and their respective 95% confidence intervals for each psychophysical fit. Non-overlapping confidence intervals indicate statistical significance of at least  $p < 0.05$ .

One subject was excluded from further analysis because of a significantly lower slope estimate in one of the three conditions so that a psychophysical threshold could not be reliably estimated from this condition.

## Results

### *Natural scene recognition is aided by natural image color in human observers*

We conducted psychophysical experiments in 10 human observers to test the effect of color presence and absence on recognition performance in a visual memory task. Recognition performance averaged across all observers is depicted in [Figure 2](#). In [Figure 2A](#), the graph plots the proportion of correct trials (symbols) and the maximum likelihood fits of a logistic function (lines) as a function of coherence for color conditions a, b, and c. Each data point corresponds to the averaged performance of 10 observers. [Figure 2B](#) plots the grand-averaged threshold estimates per stimulus condition. Comparison of grand-averaged threshold estimates evaluated at 72% correct revealed significant differences between all three color conditions. Lowest thresholds were found for the image-specific color condition (b), intermediate values for the achromatic condition (a), and highest thresholds for the colored noise

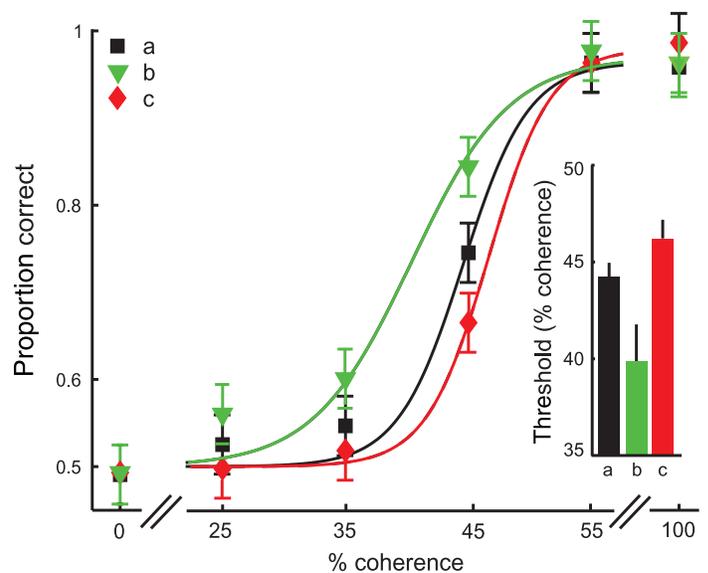


Figure 2. Behavioral performance of human observers in the recognition task. (A) Grand average ( $N = 10$ ) of proportion of correct responses plotted as a function of % coherence as well as fitted psychometric functions for each of the three color conditions a, b, and c separately. Recognition performance was best for the image-specific color condition and worst for the colored noise condition. (B) Mean threshold estimates (evaluated at 72% correct responses) as a function of color condition for the (a) achromatic condition, (b) image-specific color condition, and (c) colored noise condition. Error bars denote  $\pm 1$  standard deviation. Comparison of thresholds revealed significantly lower thresholds for the image-specific color condition when compared with the achromatic or colored noise condition. Highest thresholds were found for the colored noise condition ( $p < 0.001$  for all three comparisons based on bootstrap simulations, see Methods section for details).

condition (c). (Mean threshold estimates across subjects were (a) 44.29 with lower and upper 95th percentile confidence intervals ( $ci_{95}$ ) 41.95/45.27, (b) 39.86,  $ci_{95}$  37.7/42.05, (c) 46.22,  $ci_{95}$  45.18/48.27;  $p < 0.001$  for all comparisons based on bootstrap simulations, see Methods section for details. Psychometric functions were found to be parallel between the conditions, i.e., no significant differences in slope parameters were observed with slope estimates for (a) 0.0935,  $ci_{95}$  0.05/0.32, (b) 0.098,  $ci_{95}$  0.061/0.28, and (c) 0.063,  $ci_{95}$  0.045/0.09.) This allowed us to summarize all results at the single threshold level.

We also found this result on an individual subject level. [Figure 3](#) depicts individual threshold estimates at each of the color conditions. In 8 out of 10 subjects, lowest thresholds were found in the image-specific color condition ( $p < 0.05$ , Binomial Test). To facilitate the comparison across conditions within each subject, we inserted [Supplementary Figure 1](#), which shows individual threshold estimates for each stimulus condition in separate bar graphs per subject.

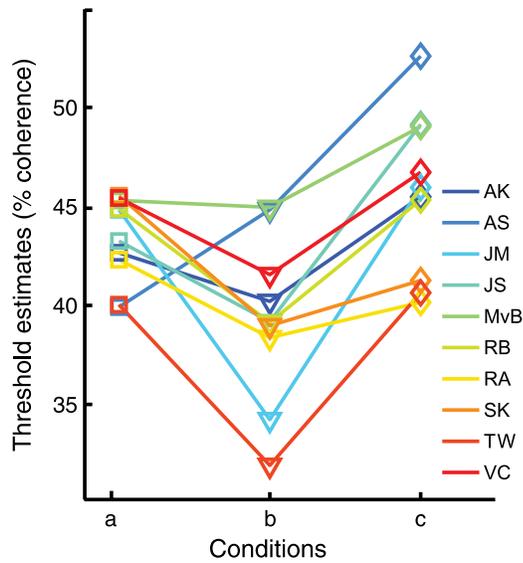


Figure 3. Individual threshold estimates for each of the 10 observers participating in the experiment. For 8 out of 10 subjects, threshold estimates were lowest in the image-specific color condition (b).

In summary, we found that in all three stimulus conditions our phase manipulation led to a monotonic reduction in recognition performance of natural images as a function of phase coherence. The interpolation of random phase noise proved to be an effective procedure to distort the spatial structure of natural images that is defined by global shapes, such as edges and contours. Our results are consistent with earlier findings on the beneficial effects of natural image color on human recognition memory for natural scenes (Gegenfurtner & Rieger, 2000; Spence et al., 2006; Wichmann et al., 2002). The present results suggest that color can lead to either enhanced or decreased recognition memory for natural scenes dependent on whether the color is diagnostic for the task at hand.

We also conducted additional experiments with four human observers involving stimulus conditions d and e. To test whether color salience could have contributed to the recognition advantage, we compared the performance for the conditions in which we interpolated between a colored noise mask and a colored image (d) to the achromatic condition (a). Figure 4A shows the recognition performance averaged across the four observers as well as the psychophysical thresholds derived at a performance level of 72% correct in Figure 4B. We found significantly lower psychophysical thresholds in condition d compared to condition a (mean threshold estimate for condition a: 44.07%,  $ci_{95}$  40.62/47.84 and condition d: 34.6,  $ci_{95}$  31.1/37.6). This was true for the average group performance as well as on a single subject level. Supplementary Figure 2 depicts single subject thresholds for these two conditions. The positive and negative effects of color on recognition

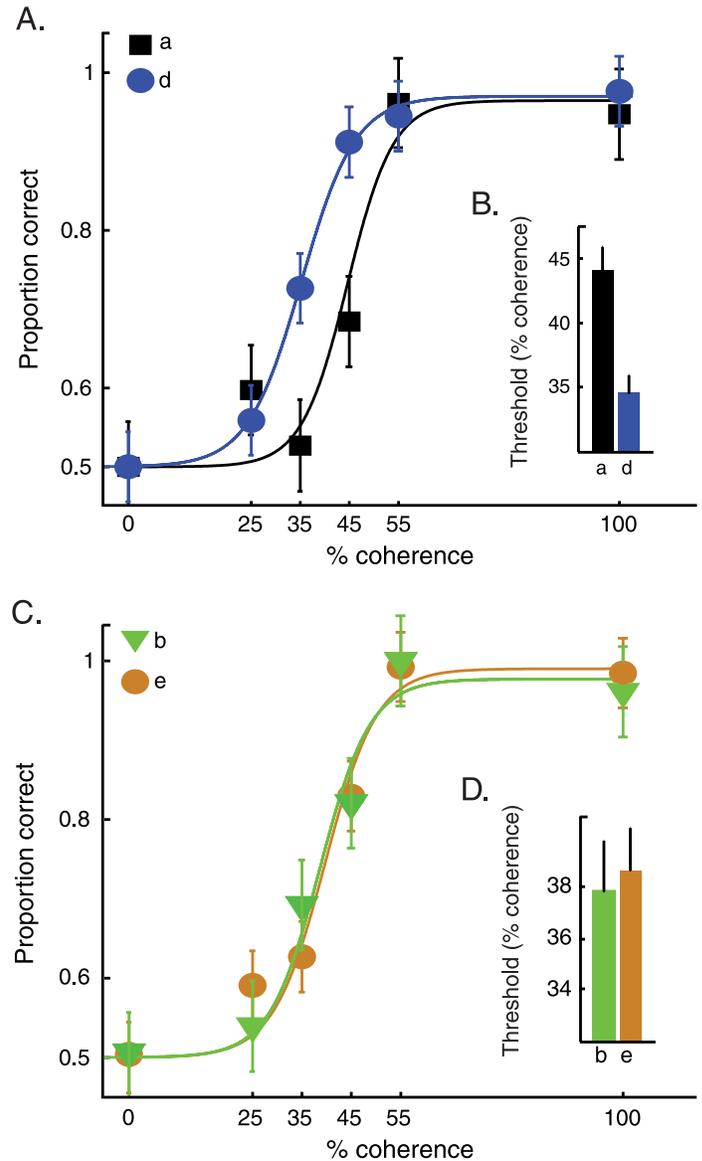


Figure 4. Pairwise comparisons of behavioral performance in human observers for conditions d and e. (A) Grand average ( $N = 4$ ) of proportion of correct responses plotted as a function of % coherence and psychometric function fitted for achromatic condition (a, black square) and color-color condition (d, blue circle). (B) Mean threshold estimates (evaluated at 72% correct responses) illustrated for the (a) achromatic condition as well as (d) color image/color noise condition. Recognition performance was significantly enhanced in condition d compared to condition e. (C) Comparison for the image-specific color condition (b, green triangle) and the color sample/achromatic test condition (e, yellow circle) showing proportion correct values along with psychometric function fits. (D) Mean threshold estimates illustrated for conditions b and e. Thresholds slightly increased for condition e compared to condition b. Error bars denote  $\pm 1$  standard deviation.

could be mediated by increased attention to either informative or non-informative regions in an image. Since these effects are equally likely to occur in condition d due to the simultaneous presence of image color and colored noise, an advantage of recognition performance in condition d over condition a likely arises from a specific, i.e., attention-independent advantage of color in recognition.

In order to assess the contribution of color at the time of retrieval, we contrasted recognition performance for conditions b and e. Both conditions are comparable regarding the color content of the sample stimulus but differ with respect to whether the test was shown in color or not. Recognition performance averaged across observers for these two conditions is depicted in Figure 4C along with the respective psychometric function fits. Figure 4D depicts the grand-averaged threshold estimates evaluated at 72% correct per stimulus condition. When comparing thresholds we found a slight threshold increment for condition e compared to b (threshold estimate for e: 38.64,  $ci_{95}$  36.44/41.53 and threshold estimate for b: 38.64%,  $ci_{95}$  33.7/41.5). Although the threshold increment was not significant on the group level, it was consistent on the individual subject level: all four subjects showed slightly enhanced thresholds for condition e when compared to condition b (see Supplementary Figure 2). Although this effect cannot be significant when comparing only four individual subjects (Binomial Test, 4 out of 4,  $p = 0.0625$ ), it might point to a small advantage for color at the retrieval stage of recognition for human observers.

## Experiments with non-human primates

### Methods

Two adult male rhesus monkeys (*Macaca mulatta*) participated in the experiments. All studies were approved by local authorities and were in full compliance with applicable guidelines (EUVD 86/609/EEC) for the care and use of laboratory animals. The behavioral task of the monkeys was a delayed matching-to-sample task very similar to the one described in the human study. The monkey was seated in front of a screen at a distance of approximately 110 cm. An initial tone indicated the potential start of a trial. The monkey initiated a trial start by grasping a lever and fixating on a small fixation spot on the center of the screen. After 1000 ms, a first stimulus appears on the screen for 250 ms, the so-called sample stimulus. The sample stimulus is presented in any of the stimulus conditions shown in Figure 1A, i.e., at any of the % coherence levels and any of the color conditions, respectively. The sample stimulus is followed by a delay period of 1500 ms during which the monkey holds

fixation. After the delay, a second stimulus, the so-called test stimulus, is presented. The test stimulus could be any of the non-degraded (100% coherence) natural images. For each color condition, the corresponding test stimulus is shown in Figure 1A on the leftmost column. The monkeys were rewarded for a lever release, whenever the test stimulus matched the sample stimulus, i.e., if the sample was either identical to the test stimulus or a degraded version (“match”). Whenever the test stimulus did not match the sample (“non-match”), the monkeys’ task was to withhold the lever release until, after a brief delay of 200 ms, a second test stimulus appeared, which always matched the sample. This procedure ensured that the monkey had to initiate a behavioral response on every trial. The monkeys were rewarded with juice for every correct trial, and randomly for the conditions in which the sample stimulus was a pure noise stimulus (i.e., 0% coherence). In each experiment, 50% of the trials were “match”, 50% were “non-match” trials, so that on pure noise trials monkeys could perform at maximally 50% correct on average. Within one session, the different trial types were randomly interleaved. Stimuli were  $7^\circ \times 7^\circ$  in size, 24-bit color depth, and presented at the center of gaze on a 21-inch monitor (ViewSonic P810) with linear luminance response as well as linear response at each color channel separately (gamma corrected). The stimulus generation procedures were identical to the ones described for the human experiment. We presented the images at four coherence levels (0%, 45%, 55%, and 100%) at each of the different color conditions.

### Sessions with novel stimuli

In each session, a set of 3–4 stimuli that were unfamiliar to the monkey was chosen. The monkeys were shown a total of 20 and 15 images for monkey N and monkey K, respectively. On average, each image was presented 84 times to the monkeys across all stimulus conditions. Thus, each image was shown about 10 times at each stimulus condition.

### Sessions with familiar stimuli

From the set of stimuli that were novel to the monkeys, three images were randomly selected and repeatedly shown in subsequent sessions. The images were repeatedly presented in a total of 17 and 19 sessions to monkeys N and K with an average rate of 300–350 trials per session for monkey N and monkey K, respectively. To ensure comparable familiarity with images across both monkeys, we only included sessions across which the performance did not show any systematic changes due to learning anymore. Therefore, we first computed the average performance for each individual session (excluding pure noise trials). We assumed that performance improved up to a certain point (i.e., session) and remained constant afterward. To find

this point, we fitted our data set with two lines. One linear fit could have a non-zero slope, to allow for systematic changes due to learning, the second linear fit was restricted to have a slope of zero, assuming no systematic change in performance. For monkey N, the minimal least-squared error (LSE,  $5.1 \times 10^{-4}$ ) was obtained when combining a fit for sessions 1 to 3 with a zero slope fit from session 4 to session 17. For monkey K, the minimal error was estimated when a linear function fit from session 1 to session 5 was combined with a zero slope linear fit from session 6 to session 19 (LSE,  $8.3 \times 10^{-4}$ ).

In order to calculate psychophysical performance, we averaged across the number of correctly identified images for each of the stimulus conditions. Because the monkeys did not work the exact same number of trials each session, we did not average across percent correct values obtained in different training sessions but across all trials the monkey performed. We subsequently derived the standard deviation from a binomial distribution. Psychometric functions (logistic function fit) were fitted using the `psignifit` toolbox (version 2.5.6 for Matlab, which implements the maximum likelihood method described by Wichmann and Hill (2001a, 2001b)). In order to do statistical comparisons between threshold estimates of different conditions, we obtained estimates for parameters of interest for psychophysical functions (slope, thresholds) through a bootstrapping procedure using the bootstrap method with BCA correction implemented in `psignifit`, based on 2000 simulations. This method is also described in detail in Wichmann and Hill (2001b). We report threshold estimates and their respective 95% confidence intervals for each psychophysical fit. Non-overlapping confidence intervals indicate statistical significance of at least  $p < 0.05$ . In the case of the threshold estimates reported in Figure 9 (comparison of conditions a vs. d and b vs. e of monkey N for familiar images), the confidence intervals overlapped. Therefore, we obtained the  $p$ -value for the statistical comparison between these conditions by bootstrapping psychometric functions. When comparing condition a to d, in 6 out of 199 samples the threshold estimate was higher in condition d than in condition a, which corresponds to a  $p$ -value of 0.03. For b vs. e, the  $p$ -value was 0.045. Since the monkeys' performance did not reach ceiling at the supra-threshold noise level (100% coherence condition), we selected the lambda parameter ("lapse rate") of the psychometric fitting function to be within a  $\pm 5\%$  range around the performance level at the 100% coherence level for each stimulus condition separately.

## Results

### **Color helps monkey observers to recognize novel as well as familiar objects in natural scenes**

To examine whether these effects might be a general characteristic of the primate visual system, we performed

psychophysical experiments in monkeys (*Macaca mulatta*) investigating the effect of color on recognition using the same stimuli and a similar task as we had used in human observers. In addition, we investigated the learning dependency of this effect by examining differences between novel and highly familiar natural images. We hypothesized that an advantage of color for novel images might lie at the level of stimulus encoding, whereas for highly familiar images color could potentially additionally act as a learned attribute of each particular object or scene. If color is an acquired feature that is bound to the representation of the stimulus in long-term memory, color might lead to a further advantage for recognition of familiar images as compared to novel images. Note that during each session, new, unfamiliar noise patterns were used for interpolation with the images. Monkeys could thus not use precise details of individual noisy stimuli but had to learn to extract structure from noisy displays, which were different every session.

In a first set of experiments, we tested the recognition performance of two adult male rhesus monkeys to novel stimuli for stimulus conditions a, b, and c. In multiple sessions, monkeys were presented with different sets of unfamiliar natural images. Figure 5 shows psychophysical performance of monkey K and monkey N for each of the three color conditions separately. Each graph plots the proportion of correct trials (symbols) and the maximum likelihood fits of a logistic psychometric function (lines) as a function of coherence. Each data point corresponds to the averaged performance during 5 and 6 sessions for monkeys K and N, respectively. Figure 5 also plots threshold estimates for the different stimulus conditions a, b, and c.

Inspection of Figure 5 reveals that recognition performance increases as a function of percent coherence in both monkeys for all color conditions, as expected. In addition, both monkeys' performances are best for the image-specific color condition, intermediate for the achromatic condition, and worst for the condition in which colored noise was added to the images. This effect can also be observed by comparing perceptual thresholds derived from the psychometric function at 72% correct performance, which usually represented the inflection point of the curves. The thresholds for each color condition are shown in Figure 5B. Comparison of the thresholds revealed significant differences between each color condition for both monkeys: recognition thresholds were lowest for the natural image condition, intermediate for the achromatic, and highest thresholds were found for the colored noise condition. We found a significant difference between all pairs of color conditions for both monkeys. For monkey K, mean thresholds are given as follows: achromatic condition (a) at 50.3%,  $ci_{95}$  47.01/52.51, image-specific color condition (b) at 47.3,  $ci_{95}$  45.59/49.3, colored noise condition (c) 56.06%,  $ci_{95}$  54.94/63.8,  $p < 0.05$  for all cases). For monkey N, mean thresholds are given as follows: (a) 50.7%,  $ci_{95}$  45.7/55.9, (b) 46.6%,  $ci_{95}$  43.8/

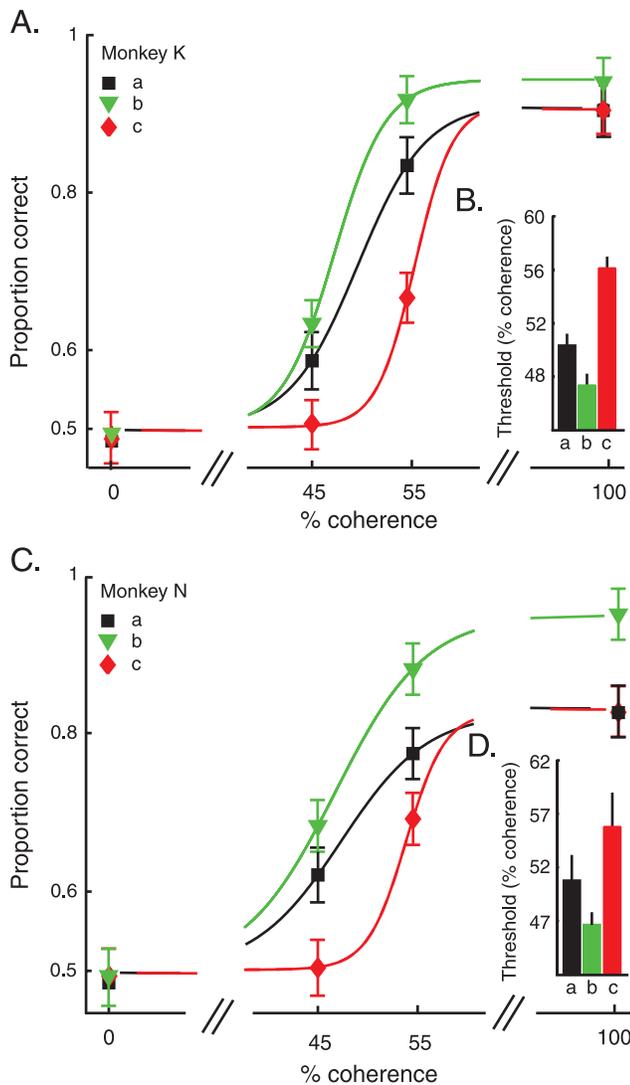


Figure 5. Behavioral performance for two monkeys in a delayed matching-to-sample task using novel stimuli. Proportion of correct responses plotted as a function of % coherence as well as fitted psychometric functions for each of the three color conditions a, b, and c separately for (A) monkey K and (B) monkey N. Different symbols/colors correspond to different color conditions: (a) achromatic condition, (b) image-specific color condition, (c) colored noise condition. Monkeys' performances were at chance in the 0% coherence condition and significantly increased at the remaining coherence levels. In addition, recognition performance also differed between the color conditions with lowest thresholds for the image-specific color condition (b). Inspection of thresholds for (B) monkey K and (D) monkey N revealed significantly lower thresholds in the image-specific color condition (b) compared to the achromatic (a) and colored noise conditions (c);  $p < 0.001$  for all three comparisons, errors bars show  $\pm 1$  standard deviation.

49.0, (c) 55.7%,  $ci_{95}$  54.5/65.98;  $p < 0.001$  for all comparisons. (As in the previous experiment, all psychometric functions were found to be parallel between the conditions, i.e., there were no significant differences in

slope parameters, see Methods section for slope estimates.)

Monkey N was also tested in conditions d and e. Figure 6 depicts recognition performance in these two conditions for novel stimuli in comparison to the achromatic condition (a, panel A) and the image-specific color condition (b, panel B), respectively. Inspection of psychophysical thresholds revealed that the monkey's performance was significantly better for condition d when compared to condition a (mean threshold estimate for condition a: 50.7%,  $ci_{95}$  45.7/55.9 and condition d: 46.68,  $ci_{95}$  44.58/49.69). When comparing condition b with condition e, we also found a significant difference between threshold estimates. The monkey's thresholds were significantly reduced in the condition in which the sample as well as the test were colored (b) compared to the condition in which only the sample stimulus appeared in color (e). (Threshold estimates for b: 46.6%,  $ci_{95}$  43.8/49.0, and condition e: 49.6,  $ci_{95}$  46.42/52.6.)

In a second set of experiments, we tested monkeys' recognition performance to familiar stimuli. Here, both monkeys were repeatedly presented with the same set of stimuli for several sessions on subsequent days. Stimuli were defined as being "familiar" to the monkeys when no systematic change in average performance across sessions was observed anymore, thereby excluding learning-related changes during the initial training sessions. Figure 7 plots

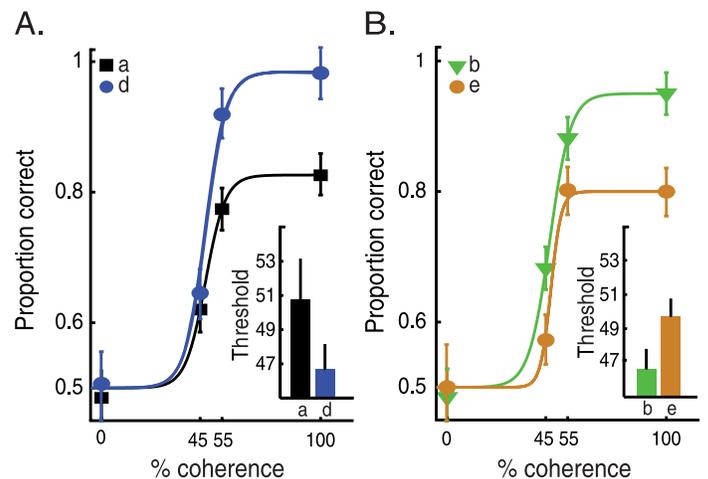


Figure 6. Pairwise comparisons of behavioral performance of one non-human primate for novel images between stimulus (A) conditions a and d and (B) conditions b and e. All plots show psychometric function fitted to proportion of correct responses and bar plots illustrate mean threshold estimates for different color conditions. Comparison of achromatic condition (a, black) to color-color condition (d, blue) reveals better recognition performance for colored images using novel images. Comparison of the image-specific color condition (b, green) to color sample/achromatic test condition (e, yellow circle) shows decreased recognition performance for condition e compared to condition b. Error bars denote  $\pm 1$  standard deviation.

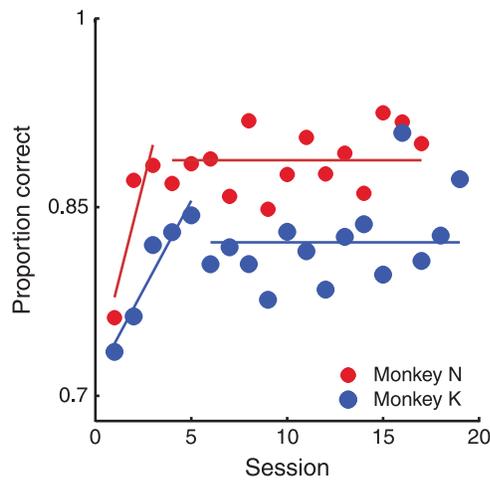


Figure 7. Average behavioral performance in recognition task across all sessions in which natural scenes were repeatedly shown to monkey K (blue) and monkey N (red). Straight lines indicate the two linear least square fits to the averaged performance across sessions yielding a minimal combined least-squared error. No systematic change in average performance could be observed for sessions 6 to 19 for monkey K and sessions 4 to 17 for monkey N. Here, different stimuli were equally familiar to both monkeys.

the averaged performance as a function of session (symbols) along with the least-error combination of two linear fits (solid lines) for monkeys K and N, respectively. While both monkeys showed improvements in recognition performance across the first few sessions, recognition performance stayed constant across sessions 6–19 for monkey K (filled circles) and sessions 4–17 for monkey N (open circles), respectively.

Figure 8 depicts the averaged behavioral performance of both monkeys for the sessions for which behavioral performance had reached an asymptote. Each graph plots the proportion of correct trials (symbols) and the maximum likelihood fits of a logistic psychometric function (lines) as a function of % coherence separately for each color condition for monkey N (upper panel) and monkey K (lower panel).

The pattern of results obtained here are comparable to the pattern of results obtained in experiments using novel stimuli: although recognition performance as measured by threshold estimates is significantly decreased for familiar stimuli (see discussion of learning-related changes below), the pattern of results for the different color conditions was the same as for novel images; recognition performance was significantly increased by image-specific color and significantly reduced in the colored noise condition. Threshold estimates for monkey K are given as follows: (a) at 47.4%,  $ci_{95}$  46.6/48, (b) 44.6,  $ci_{95}$  44.04/45.3, (c) 52.3%,  $ci_{95}$  51.6/53.05,  $p < 0.05$  for all cases; for monkey N: (a) 42.6%,  $ci_{95}$  40.5/44.7, (b) 39.72%,  $ci_{95}$  37.74/41.36, (c) 45.8%,  $ci_{95}$  45/46.6;  $p < 0.001$  for all comparisons. For both monkeys, no significant differences

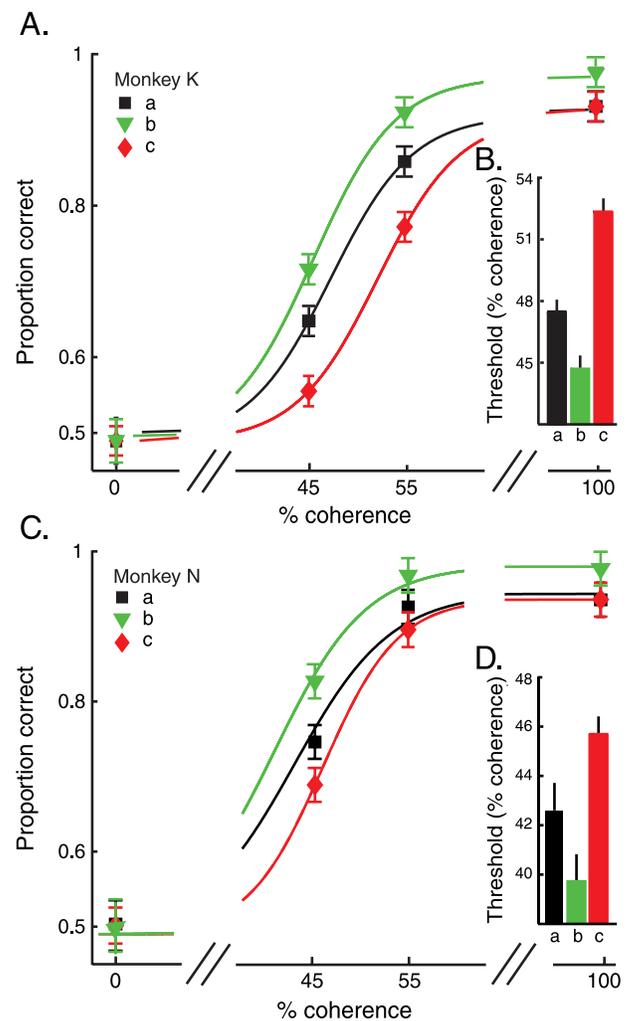


Figure 8. Behavioral performance for familiar stimuli. Proportion of correct responses plotted as a function of % coherence as well as fitted psychometric functions for each of the three color conditions separately for (A) monkey K and (C) monkey N. Different symbols and line styles correspond to different color conditions and are equivalent to denotations used in Figure 5. As seen for novel stimuli, monkeys' performances were at chance in the 0% coherence condition and significantly increased at the remaining coherence levels ( $p < 0.001$  for all comparisons and both monkeys). In addition, recognition performance also differed between the color conditions. Inspection of perceptual thresholds, for (B) monkey K and (D) monkey N, revealed significantly lower thresholds in the image-specific color condition (b) compared to the achromatic (a) and colored noise conditions (c);  $p < 0.001$  for all three comparisons, error bars show standard deviation.

in slope estimates were found between the color conditions, indicating only a parallel shift of psychometric functions between the conditions, i.e., as for human observers we were able to summarize our results at the single threshold level (see Methods section for slope estimates).

Equivalent to the experiments involving novel stimuli, monkey N was also tested in conditions d and e using familiar images. Figure 9 shows psychophysical functions

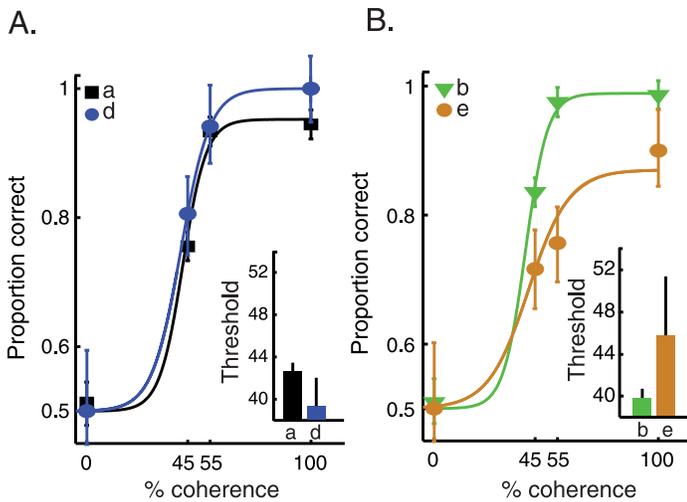


Figure 9. Pairwise comparisons of behavioral performance of one non-human primate for familiar images between stimulus (A) conditions a and d and (B) conditions b and e. All plots show psychometric function fitted to proportion of correct responses and bar plots illustrate mean threshold estimates for different color conditions. Comparison of achromatic condition (a, black) to color–color condition (d, blue) reveals better recognition performance for colored images using novel images. Comparison of the image-specific color condition (b, green) to color sample/achromatic test condition (e, yellow circle) shows decreased recognition performance for condition e compared to condition b. Error bars denote  $\pm 1$  standard deviation.

as well as threshold estimates equivalent to Figure 6. In both conditions, the monkey's performance as measured by psychophysical thresholds was enhanced for familiar compared to novel images. In addition, we obtained matching results using familiar images regarding the comparison between the different color conditions: the monkey's performance was significantly better for condition d when compared to condition a (threshold estimates for a: 42.6%,  $ci_{95}$  40.5/44.7 and d: 39.3%,  $ci_{95}$  16.61/44.1,  $p = 0.03$ , see Methods section for details). Furthermore, we also found lower recognition performance for stimulus condition e compared to condition b resulting in a significant increase in threshold estimates for condition e compared to condition b (threshold estimates for b: 39.72%,  $ci_{95}$  37.74/41.36, and condition e: 45.8,  $ci_{95}$  35.94/57.23,  $p = 0.045$ , see Methods section for details).

### **The beneficial effect of color on recognition is comparable for novel and highly familiar natural images**

Both monkeys showed substantial improvements in performance for familiar images when compared to novel images. At each level of coherence, behavioral performance was enhanced for both monkeys for familiar as

compared to novel images, which resulted in significantly decreased psychophysical thresholds across all color conditions for familiar images in both monkeys (monkey K thresholds novel: 50.3, 47.3, and 56.6 for conditions a, b, and c, respectively; thresholds familiar: 44.6, 47.4, and 52.3; monkey N thresholds novel: 50.7, 46.6, 55.7, 46.68, and 49.6 for conditions a, b, c, d, and e, respectively, thresholds familiar: 39.7, 42.6, 45.8, 39.3, and 45.8). When comparing the thresholds between the color conditions for novel and familiar stimuli, the pattern of differences was the same for conditions a, b, and c in both monkeys (i.e., lowest thresholds for image-specific condition, highest thresholds for colored noise condition). This was also true when comparing stimulus conditions a and b to stimulus conditions d and e in monkey N. For both novel and familiar stimuli, we found an advantage of the colored condition (d) compared to the achromatic condition, as well as an advantage for the condition in which both sample and test stimuli were colored (b) compared to the condition in which the test stimulus was achromatic (e). In addition, the magnitude of the beneficial effect of natural image color was comparable for novel as well as highly familiar stimuli with respect to the change in thresholds (% change in thresholds, image-specific color vs. achromatic color; for novel images, monkey K: 6.4%, monkey N: 8.7% and familiar images, monkey K: 6.3%, monkey N: 6.5%). Taken together, this suggests that learning improves performance non-specifically across all color conditions but does not lead to further specific performance advantages or detriments related to the color manipulations. Learning thus seems to generally allow the monkeys to extract luminance- and color-related signals from the noisy displays more efficiently.

In summary, we found a beneficial effect of image-specific color on object recognition in both monkeys. Since our image degradation procedure allowed for independent modulation of color content, our results suggest that color, separate of luminance-based shape information, helped monkeys recognize natural images. These results demonstrate that natural image color is beneficial for visual memory performance involving natural scenes in the non-human primate. In addition, both monkeys showed substantial improvements in recognition performance for familiar stimuli. For familiar stimuli, the effect of color on recognition performance was comparable to that of novel images. This result implies that color might mainly play a role in perceptual and short-term memory processing, rather than long-term memory processes.

## **Discussion**

Our results show a 6–8% coherence-related improvement in recognition performance for noise-degraded

colored natural images when compared to achromatic natural images. In addition, the addition of color to the noise used for degradation impaired recognition performance when compared to the use of achromatic noise. Color can thus help or hinder the extraction of task relevant information from visual displays, depending whether it is congruent (image-specific color condition) or incongruent (in form of colored noise) with luminance-based structure information. These effects were present in human subjects as well as in each of two adult rhesus monkeys that were tested with an identical visual recognition paradigm.

Our results confirm findings from earlier studies showing a beneficial effect of color on recognition performance for natural scenes. Two possible mechanisms have been mainly suggested to underlie the advantage of image color for recognition (Gegenfurtner & Rieger, 2000; Spence et al., 2006; Wichmann et al., 2002). First, during the encoding phase, i.e., when the observer initially perceives the image, color is likely to improve image segmentation. Segmentation refers to the process of segregating a complex scene into its constituent regions, surfaces, and objects. In images, color helps define spatial contours, surfaces, and boundaries, irrespective of what the exact color of the object is. Due to a better segmentation, more information about the scene and its objects is encoded and can subsequently be used for recognition. In naturally colored images, color can thus provide an additional segmentation cue. In contrast, in achromatic images color could not be used for segmentation and in images that were degraded by colored noise meaningful scene segmentation might even be obstructed by color. An additional explanation for why colored noise is more effective in worsening performance than achromatic noise is that it might be easier to “see” or “interpret” shapes or objects in a colored rather than an achromatic setting, even if the shapes are only noise. The subjective interpretation of the scene might subsequently interfere with the actual object that is embedded in the noise. In contrast to the hypothesis that colored noise would obstruct scene perception to a greater extent than achromatic noise, one might argue the contrary: scene segmentation could improve in the case of colored noise because hue provides additional information to distinguish noise from image features. However, our results suggest that this is likely not the case.

An alternative explanation for the pattern of results for conditions a, b, and c could lie in the fact that color is a very salient visual feature that can engage visual attention (Davidoff, 1991). By drawing the observers’ attention to either informative or uninformative regions in an image, one might expect that if color is informative (as in the image-specific condition, b) it helps recognition and if it is uninformative (as in the colored noise condition, c) color hinders recognition. Therefore, an important control condition is stimulus condition d, in which a colored noise mask is interpolated with a colored image. For this condition, recognition performance was still better than in the achromatic condition. This was true in both human

and monkey observers. Since possible attention-related effects are equally beneficial and distractive in condition b, a superiority of this condition compared to the achromatic condition suggests that color saliency cannot be the sole reason for a color advantage in recognition. Taken together, our findings thus support claims from previous studies in human observers (Wichmann et al., 2002) and imply an attention-independent effect of color for recognition. As we have already argued, this could possibly lie in enhancing image segmentation.

The role of color for segmentation might be particularly important in cases in which contours and regions are poorly defined by variations in luminance alone. This situation is well reflected in our naturally colored images (condition b) at the intermediate degradation levels. As we added noise, contour information based on luminance alone became less reliable. The same was true for condition d involving colored noise. Here, contour information became less reliable as colored noise was added to the scene. Our findings are in accordance with previous studies that have emphasized the important role of image segmentation under conditions of degraded shape cues (Li & Lennie, 2001; Mollon, 1989; Shevell & Kingdom, 2008). Given that our stimuli consisted of objects embedded in natural scenes, improved segmentation is likely to have contributed to advantages in recognition performance. On the other hand, image segmentation would probably not be a mechanism used for the recognition of isolated objects, since they do not have to be segregated from the background. This is line with an earlier study that reported no effect of color on recognition performance for single objects (Joseph & Proffitt, 1996). Further support for the image-segmentation-by-color hypothesis comes from psychophysical results showing the use of color in natural scene segmentation (Fine et al., 2003), reports on multiple spatially tuned chromatic mechanisms that can serve image segmentation in low-level perceptual tasks (Gegenfurtner & Kiper, 1992; Hansen & Gegenfurtner, 2006; Li & Lennie, 2001) as well as electrophysiological and psychophysical studies illustrating the importance of color for the analysis of visual form (De Valois & Switkes, 1983; Johnson, Hawken, & Shapley, 2001).

Secondly, it has been proposed that color can act as an additional retrieval cue at recognition (Gegenfurtner & Rieger, 2000; Spence et al., 2006; Wichmann et al., 2002). This could be the result of what is commonly known as the encoding specificity principle (Tulving, 1972), where successful retrieval from memory is dependent upon the overlap of information at the time of encoding and retrieval. In all our conditions, the overlap of shape information between encoding and retrieval is dependant on the noise level. The more object shape is degraded, the less it can function as a retrieval cue for subsequent recognition. Shape degradation is equally strong in the black and white and color conditions. Thus, shape cues were equally informative (or non-informative) at the time

of retrieval. For naturally colored images, the particular color of an object can act as an additional cue helping recognition since the same color information is present at the encoding as well as the retrieval phase. However, this is not the case for achromatic images degraded with colored noise: color that was present in the sample stimulus could not be found in the achromatic test stimulus and was therefore useless. Thus, colored noise interfered with recognition in a two-fold way: it disturbed a meaningful segmentation process during the encoding and was useless as a possible retrieval cue at recognition.

Another indication for a possible usage of color as a retrieval cue comes from the slight recognition advantage for stimulus conditions b and d over condition e. In both conditions b and e, color was present at the time of encoding as well as at the time of retrieval. In contrast, in condition d color could only help encoding, since the relevant test stimulus was always shown in its achromatic version. Although the size of this effect appeared to be minor compared to the encoding-related improvement in performance due to color, we suggest that in order for color to exert its benefits on recognition processes, a sensible color–shape association is required during the encoding phase as well as at the time of retrieval.

Some color–shape associations comprise conceptual knowledge about the object. In the case of color, this phenomenon is termed “color diagnosticity” (Tanaka & Presnell, 1999). “Color diagnosticity” refers to the association between an object and a particular color. A canonical example is a yellow banana, where the color yellow is an intrinsic mnemonic property of the object banana. Man-made objects as well as objects emerging in multiple or various colors are thought to be low in diagnosticity or not diagnostic at all. We would therefore refer to flowers, parrots, and butterflies as non-diagnostic images, because there is not one single characteristic color these objects are associated with but they rather appear either in many colors simultaneously (parrot, butterfly) or in a range of colors (flowers). It can be argued that monkey images might be somewhat diagnostic in color since they appear in similar color gamut such as black and brown. However, all images exhibited a range of different colors for the monkeys as well as their surroundings and background.

Studies on intrinsic object properties have suggested that when shape is less diagnostic or degraded, diagnostic properties such as color become more heavily weighted on recognition judgments (Naor-Raz et al., 2003). Since we used images and objects with no or low color diagnosticity, this mechanism is less likely to have contributed to our results.

Our findings demonstrate beneficial effects of natural image color for visual memory processing in the non-human primate. In the past, only a few studies have used natural scenes to study the influence of color on visual perception in the macaque monkey (Delorme, Richard, & Fabre-Thorpe, 2000; Vogel, 1999). Both of these studies

did not find a considerable effect of color for either task performance or reaction time in non-human primates (as well as human observers), which is in contrast to our findings. Several factors might have contributed to the observed discrepancy. Firstly, in both above-mentioned studies, all monkeys showed very high performance levels (>90% correct responses), most likely due to the fact that the images were not degraded. Some authors have argued that color might be particularly useful when the shape of objects or scenes does not provide sufficient information for recognition (Biederman & Ju, 1988; Tanaka et al., 2001). In our study, we could assess the effect of image color under shape-degraded conditions and indeed found an advantage for recognition performance. Secondly, the experimental paradigms employed were considerably different. In the two aforementioned studies, the monkeys’ task was a categorization task, i.e., monkeys had to detect a target category, for example a tree or an animal and respond as fast as possible with an associated saccade movement to a particular direction. Both studies focused on rapid categorization of predefined object categories. Our paradigm incorporated a short-term memory process, during which a mnemonic representation of the visual input has to be compared with a following test stimulus. In a categorization task, color could have only contributed during the encoding of the stimulus. In contrast, in a recognition paradigm, color can also act as an additional retrieval cue. Therefore, our results are likely to not only reflect the advantage of color for encoding but also for retrieval processes. Thirdly, in our study natural images differed with respect to their image content. There was no predefined and trained target category that had to be detected or differentiated. If objects that belong to different categories are less similar regarding their shape than their color, i.e., if color is not a discriminative feature between the two categories, it might be more useful to rely on shape cues instead of color cues to solve the task. Using a variety of randomly selected images makes such a strategy less likely, since natural images show higher variability with respect to the shapes and colors they contain.

## Conclusion

In order to reliably link the results gathered in human observers to recognition performance in rhesus monkeys, we employed the identical visual memory paradigm in both species. In both experiments, we found a comparable benefit of natural image color for visual memory performance as well as a comparable detrimental effect of colored noise on recognition. These results extend previous reports about similarities between the two species in color vision at early stages of processing to the domain of high-level vision and visual memory. Our

behavioral evidence suggests that the mechanisms by which color promotes or interferes with recognition could be identical in both species: Color helps perceptual encoding and retrieval from short-term memory. In human observers, color can also promote recognition memory through the association of a particular object with a particular color in diagnostic objects (Rossion & Pourtois, 2004; Tanaka et al., 2001). Although many non-human animals are thought to have a semantic memory system (Tulving, 2002), it is not clear whether rhesus monkeys will benefit from this association to the same extent as humans. This is particularly questionable for novel objects. However, monkeys show the beneficial effect of color also for novel objects, strongly suggesting that their recognition performance is possibly based on matching the sensory attributes of images irrespective of a conceptual association between a particular color and an object. Another aspect supporting this notion is that we did not find any interaction between the effect of color and learning. One might argue that if the relationship between a particular color and object through learning processes had been formed, an additional benefit of color for the recognition of familiar stimuli would have been expected. However, we studied learning only on the time scale of days. Therefore, we cannot rule out that a possible interaction of learning and the effect of color on recognition might exist for learning over a much longer time scale.

Finally, it has been argued that the evolutionary advantage of color vision in the primate may lie preferentially in the detection and discrimination of edible fruits in a complex visual environment, such as a forest (Sumner & Mollon, 2000; Mollon, 1989; Regan et al. 2001). While this is a plausible hypothesis, our findings show an advantage of natural image color for a much broader scope of natural visual stimuli. Thus, it could be the case that the evolutionary advantage is by no means limited to the recognition of food sources. Rather, our results infer that color helps primates to recognize and differentiate the entire spectrum of behaviorally relevant objects.

## Acknowledgments

This research was supported by the Max Planck Society, DFG SFB550 and RA1025-1/2. The authors thank Felix Wichmann, Jakob Macke, and Andreas Bartels for helpful comments and suggestions on the manuscript and throughout the project.

Commercial relationships: none.  
Corresponding author: Stefanie Liebe.  
Email: sliebe@tuebingen.mpg.de.

Address: Max Planck Institute for Biological Cybernetics, Department of Physiology of Cognitive Processes, Spemannstrasse 38, 72076 Tübingen, Germany.

## References

- Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, *20*, 38–64. [PubMed]
- Clifford, C. W., Holcombe, A. O., & Pearson, J. (2004). Rapid global form binding with loss of associated colors. *Journal of Vision*, *4*(12):8, 1090–1101, <http://journalofvision.org/4/12/8/>, doi:10.1167/4.12.8. [PubMed] [Article]
- Croner, L. J., & Albright, T. D. (1999). Segmentation by color influences responses of motion-sensitive neurons in the cortical middle temporal visual area. *Journal of Neuroscience*, *19*, 3935–3951. [PubMed] [Article]
- Dakin, S. C., Hess, R. F., Ledgeway, T., & Achtman, R. L. (2002). What causes non-monotonic tuning of fMRI response to noisy images? *Current Biology*, *12*, R476–R477. [PubMed] [Article]
- Davidoff, J. B. (1991). *Cognition through colour*. Cambridge, MA: MIT Press.
- Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: A study in monkeys and humans. *Vision Research*, *40*, 2187–2200. [PubMed]
- De Valois, K. K., & Switkes, E. (1983). Simultaneous masking interactions between chromatic and luminance gratings. *Journal of the Optical Society of America*, *73*, 11–18. [PubMed]
- Fine, I., MacLeod, D. I., & Boynton, G. M. (2003). Surface segmentation based on the luminance and color statistics of natural scenes. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *20*, 1283–1291. [PubMed]
- Gegenfurtner, K. R., & Kiper, D. C. (1992). Contrast detection in luminance and chromatic noise. *Journal of the Optical Society of America A, Optics and Image Science*, *9*, 1880–1888. [PubMed]
- Gegenfurtner, K. R., & Kiper, D. C. (2003). Color vision. *Annual Review in Neuroscience*, *26*, 181–206. [PubMed]
- Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, *10*, 805–808. [PubMed] [Article]
- Hanna, A., & Remington, R. (1996). The representation of color and form in long-term memory. *Memory & Cognition*, *24*, 322–330. [PubMed]
- Hansen, T., & Gegenfurtner, K. R. (2006). Higher level chromatic mechanisms for image segmentation. *Journal of Vision*, *6*(3):5, 239–259, <http://journalofvision.org/6/3/5/>, doi:10.1167/6.3.5. [PubMed] [Article]

- Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, *9*, 1367–1368. [[PubMed](#)]
- Jacobs, G. H. (1996). Primate photopigments and primate color vision. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 577–581. [[PubMed](#)] [[Article](#)]
- Johnson, E. N., Hawken, M. J., & Shapley, R. (2001). The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nature Neuroscience*, *4*, 409–416. [[PubMed](#)]
- Joseph, J. E., & Proffitt, D. R. (1996). Semantic versus perceptual influences of color in object recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 407–429. [[PubMed](#)]
- Koida, K., & Komatsu, H. (2007). Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nature Neuroscience*, *10*, 108–116. [[PubMed](#)]
- Kusunoki, M., Moutoussis, K., & Zeki, S. (2006). Effect of background colors on the tuning of color-selective cells in monkey area V4. *Journal of Neurophysiology*, *95*, 3047–3059. [[PubMed](#)] [[Article](#)]
- Li, A., & Lennie, P. (2001). Importance of color in the segmentation of variegated surfaces. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *18*, 1240–1251. [[PubMed](#)]
- Mollon, J. D. (1989). “Tho’ she kneel’d in that place where they grew...” The uses and origins of primate colour vision. *Journal of Experimental Biology*, *146*, 21–38. [[PubMed](#)] [[Article](#)]
- Mullen, K. T., Beaudot, W. H., & McIlhagga, W. H. (2000). Contour integration in color vision: A common process for the blue–yellow, red–green and luminance mechanisms? *Vision Research*, *40*, 639–655. [[PubMed](#)]
- Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is color an intrinsic property of object representation? *Perception*, *32*, 667–680. [[PubMed](#)]
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, *41*, 176–210. [[PubMed](#)]
- Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, *69*, 529–541.
- Ostergaard, A. L., & Davidoff, J. B. (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 579–587. [[PubMed](#)]
- Piotrowski, L. N., & Campbell, F. W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception*, *11*, 337–346. [[PubMed](#)]
- Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *41*, 797–827. [[PubMed](#)]
- Rainer, G., Lee, H., & Logothetis, N. K. (2004). The effect of learning on the function of monkey extrastriate visual cortex. *Plos Biology*, *2*, E44. [[PubMed](#)] [[Article](#)]
- Rainer, G., & Miller, E. K. (2000). Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron*, *27*, 179–189. [[PubMed](#)] [[Article](#)]
- Regan, B. C., Julliot, C., Simmen, C., Viénot, F., Charles-Dominique, P., & Mollon, J. D. (2001). Fruits, foliage and the evolution of primate colour vision. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *356*, 229–283. [[PubMed](#)] [[Article](#)]
- 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. [[PubMed](#)]
- Shevell, S. K., & Kingdom, F. A. (2008). Color in complex scenes. *Annual Review of Psychology*, *59*, 143–166. [[PubMed](#)]
- Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, *15*, 161–171. [[PubMed](#)]
- Spence, I., Wong, P., Rusan, M., & Rastegar, N. (2006). How color enhances visual memory for natural scenes. *Psychological Science*, *17*, 1–6. [[PubMed](#)]
- Sumner, P., & Mollon, J. D. (2000). Catarrhine photopigments are optimized for detecting targets against a foliage background. *Journal of Experimental Biology*, *203*, 1963–1986. [[PubMed](#)] [[Article](#)]
- Tanaka, J., Weiskopf, D., & Williams, P. (2001). The role of color in high-level vision. *Trends in Cognitive Sciences*, *5*, 211–215. [[PubMed](#)]
- Tanaka, J. W., & Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics*, *61*, 1140–1153. [[PubMed](#)]
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25. [[PubMed](#)]

- Vogels, R. (1999). Categorization of complex visual images by rhesus monkeys. Part 1: Behavioural study. *European Journal of Neuroscience*, *11*, 1223–1238. [[PubMed](#)]
- Wichmann, F. A., Braun, D. I., & Gegenfurtner, K. R. (2006). Phase noise and the classification of natural images. *Vision Research*, *46*, 1520–1529. [[PubMed](#)]
- Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*, 1293–1313. [[PubMed](#)] [[Article](#)]
- Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*, 1314–1329. [[PubMed](#)] [[Article](#)]
- 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. [[PubMed](#)]
- 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 and Performance*, *19*, 899–911. [[PubMed](#)]
- Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, *31*, 995–1003. [[PubMed](#)]