Research Article

Color Naming, Lens Aging, and Grue

What the Optics of the Aging Eye Can Teach Us About Color Language

Joseph L. Hardy,1,2 Christina M. Frederick,3 Paul Kay,4 and John S. Werner1

1Department of Ophthalmology and Section of Neurobiology, Physiology and Behavior, University of California, Davis Medical Center;2Posit Science, San Francisco, California;3Department of Psychology, University of California, Berkeley; and 4International Computer Science Institute, Berkeley, California

ABSTRACT—Many languages without separate terms for green and blue are or were spoken in locations receiving above-average exposure to ultraviolet-B (UV-B) radiation. It has been proposed that this correlation is caused by premature lens aging. This conclusion was supported by an experiment in which younger observers used the term “blue” less often when they described simulated paint chips filtered through the equivalent of an older observer’s lens—removing much short-wavelength light—than when they described the unfiltered versions of the same paint chips. Some stimuli that were called “blue” without simulated aging were called “green” when filtered. However, in the experiment reported here, we found that the proportion of “blue” color-name responses did not differ between younger subjects and older observers with known ocular media optical densities. Color naming for stimuli that were nominally green, blue-green, or blue was virtually identical for older and younger observers who viewed the same (unfiltered) stimuli. Our results are inconsistent with the lens-brunescence hypothesis.

Most humans can discriminate millions of colors, yet in every language studied systematically, the number of basic color terms (BCTs) is comparatively small. In English, for example, there are 11 BCTs, that is, black, white, red, yellow, green, blue, purple, brown, orange, pink, and gray (Berlin & Kay, 1969). BCTs refer to the smallest collection of simple words with which a speaker can name any color. A few languages have more than 11 BCTs—Russian has distinct BCTs for light blue and dark blue, Hungarian for light red and dark red—but many languages have fewer. Welsh and most other Celtic languages, for example, do not have distinct BCTs for green and blue (Lazar-Meyn, 2004), and the same is true for almost all unwritten languages. The set of color terms within a language is not random. Rather, there is a distinct tendency for languages to evolve color terms in a particular order. For example, no language is known to have distinct words for blue and green and yet fail to distinguish red and yellow. The origins and mechanisms underlying the regularity of color naming across languages have been topics of considerable investigation (e.g., Heider, 1972; Kay & Maffi, 1999; Kay & McDaniel, 1978) and debate (e.g., Davidoff, Davies, & Robertson, 1999; Hickerson, 1971; Lucy, 1997).

Recently, Lindsey and Brown (2002) have offered an interesting hypothesis to explain why some languages lack BCTs that distinguish green and blue. According to their argument, most so-called grue languages (those with a single term covering both green and blue) occur in geographic locations closer to the equator or at higher elevations than the locations of languages that have separate terms for green and blue. According to their argument, most grue languages tend to receive above-average levels of ultraviolet-B (UV-B) radiation from the sun. High UV-B exposure is linked to accelerated ocular media aging, most particularly of the crystalline lens (Javitt & Taylor, 1994–1995; Werner, Peterzell, & Scheetz, 1990; Young, 1991). As the crystalline lens ages, a process known as brunescence occurs. The lens becomes denser and more opaque, allowing less light, especially at shorter wavelengths, to reach the retina (Weale, 1988). Individuals experiencing premature lens aging would receive less short-wavelength light at the retina when viewing the same

Address correspondence to Joseph L. Hardy, Department of Ophthalmology, 4860 Y St., Suite 2400, Sacramento, CA 95817; e-mail: joe.hardy@positscience.com.
stimuli as people with more transparent lenses. Lindsey and Brown argued that this reduced exposure to short-wavelength light would reduce the need for the color term “blue,” usually assigned to stimuli predominantly composed of energy at shorter visible wavelengths. Thus, they proposed that languages spoken by cultures located in high-UV-B areas would tend to be grue languages.

To test this hypothesis, Lindsey and Brown (2002) performed a color-naming experiment in which they simulated the effects of lens brunescence in young individuals. This simulation took advantage of the lens-aging model of Pokorny, Smith, and Lutze (1987) to transform stimulus chromaticities to simulate the effects of optical aging. Young observers were shown simulated Munsell chips (a set of color standards) made yellower and darker in the exact proportions prescribed by the Pokorny et al. lens model to simulate the effects of age-related lens opacity changes of individuals ages 30 to 100. As predicted by Lindsey and Brown’s lens-brunescence hypothesis, observers used the term “blue” progressively less as the simulated lens was increasingly aged. Chips previously called “blue” were called “green” or “gray” after being transformed by simulated aging.

Lindsey and Brown concluded that the results of this experiment show that premature lens brunescence could lead to a reduced need for the term “blue.” This argument has created considerable interest and some controversy (e.g., Lazar-Meyn, 2004; Lindsey & Brown, 2004; Regier & Kay, 2004).

Although Lindsey and Brown’s (2002) aging simulation was mathematically consistent with a valid model of lens brunescence, the effects of lens aging can be tested more directly by testing individuals of various chronological ages and, thus, a range of ocular media optical densities (ODs). The results of this more direct test could differ from the results of a simulation. Younger observers viewing stimuli designed to simulate the optics of the aging eye experience lower light levels and higher proportions of long-wavelength light for a matter of minutes. Older individuals with brunescent lenses live with their optics continuously, as do people with more brunescent lenses as a result of living in high-UV-B areas. Although Lindsey and Brown did have observers briefly (3 min) adapt to the background color prior to the experiment, chromatic adaptation can operate on very long time scales, on the order of days (Neitz, Carroll, Yamauchi, Neitz, & Williams, 2002) or even months (Delahunt, Webster, Ma, & Werner, 2004). Chromatic adaptation has the effect of changing how the visual system interprets the light reaching each of the three photoreceptor types (Jameson & Hurvich, 1956; Uchikawa, Uchikawa, & Boynton, 1989). Long-term chromatic adaptation could largely compensate for changes in the average spectral distribution of light reaching the retina due to age-related changes in the lens. Thus, although such a simulation can closely replicate the patterns of wavelengths reaching the retina in observers with differing ocular media ODs, it is unlikely to accurately simulate the perceptual experience of observers with naturally brunescent lenses.

To test Lindsey and Brown’s (2002) lens-brunescence hypothesis, we compared color naming in individuals with a range of known ocular media ODs. We did this in two ways. In one set of conditions, we tested color naming in groups of younger and older observers with the same standard stimulus set used by Lindsey and Brown. In a separate set of conditions, we tested color naming in each of these groups using stimuli that simulated the effects of ocular media present in the complementary age group (i.e., younger observers viewed stimuli filtered through the simulated ocular media of older observers, and older observers viewed stimuli filtered through the simulated ocular media of younger observers).

If Lindsey and Brown’s (2002) proposal is correct, color naming by older observers in the standard condition would be expected to be similar to color naming by younger observers in the simulated-aging condition. Also, color naming by younger observers in the standard condition would be expected to be similar to color naming by older observers in the simulated-“youthening” condition. Specifically, stimuli denoted in the Munsell color system as “blue-green” and “blue” would be expected to be identified with the color name green more often by older subjects viewing standard stimuli and younger subjects viewing simulated-aging stimuli than by younger subjects viewing standard stimuli and older subjects in the simulated-youthening condition. Additionally, in the standard stimulus conditions, a strong negative correlation between ocular media OD and use of the color term blue would be expected.

Evidence from studies examining the effects of aging on unique hue loci (Schefrin & Werner, 1990), hue scaling (Schefrin & Werner, 1993), and color naming in Japanese (Okajima, Yamashita, Takamura, Watanabe, & Tsuchiya, 2002), however, casts doubt on these predictions. These studies show that older observers, despite their increased ocular media OD, tend to use linguistic color descriptors in much the same way as younger observers. This suggests that when older observers and younger observers are presented with the same stimuli, their color naming should be similar.

**METHOD**

**Observers**

All observers were born in the United States and were native English speakers. They were provided complete ophthalmic and optometric examinations prior to inclusion in this study. Only visually healthy (e.g., no clinically significant cataracts or retinal disorders) observers were considered for this experiment. No observers had a history of cataract surgery, and, thus, all observers had intact crystalline lenses. This study included a total of 20 observers in two age groups: 10 observers in the younger group (mean = 23.2 years of age, range = 18–29) and 10 in the older group (mean = 73.9 years of age, range = 68–79). The ratio of males to females in both groups was 1:1.

Written informed consent, based on a protocol approved by the...
Ocular Media Measurement

We determined the ocular media OD for one eye of each individual before he or she participated in the color-naming portion of this experiment. Ocular media OD was estimated based on the individual’s scotopic (low-light-level) spectral sensitivity, using a variation of a technique described by Norren and Vos (1974). The rationale for this method is that under scotopic conditions, relative spectral sensitivity is dependent on the shape of the rhodopsin (rod photopigment) absorption spectrum and the ocular media OD spectrum. The absorption spectrum of rhodopsin is essentially invariant between individuals. Thus, any differences between observers in relative sensitivity to various wavelengths of light under scotopic conditions are due to differences in ocular media OD. Differences in scotopic spectral sensitivity between the individual observer and the Commission Internationale de l’Eclairage (CIE) standard observer were fitted with varying amounts of ocular media OD in proportions prescribed by the lens-aging model of Pokorny et al. (1987). The function that provided the best least-squares fit to these data was taken as the individual’s ocular media OD (see Fig. 1).

To measure the scotopic spectral sensitivity function, we used a Maxwellian-view optical system that permitted monochromatic light (seven wavelengths from 410 through 600 nm) to be imaged on the retina as an annulus (inner diameter of 7° and outer diameter of 15°) flickering as a 3-Hz square wave at 100% modulation. The observer’s task was to adjust stimulus intensity by varying a neutral-density wedge with a potentiometer until the stimulus was just detected. Following 30 min of dark adaptation, and a minimum of three practice trials, observers made a series of three threshold settings for each wavelength, presented in random order. Sensitivity was defined by the reciprocal of the energy at the geometric mean setting.

Color-Naming Experiment

In the standard stimulus condition, stimuli were colorimetrically simulated versions of the 40 Value-6 Munsell chips used in the 1997 World Color Survey (Kay, Berlin, Maffi, & Merrifield, 1997). Stimuli were presented on a CRT. The chromaticities used were from Newhall, Nickerson, and Judd (1943), who made colorimetric measurements of these chips under CIE illuminant C.

In the simulated-aging and simulated-youthening conditions, these stimuli were adjusted to simulate the effects of the ocular media OD of an average observer in the complementary age group. In the simulated-aging condition, the younger individual’s ocular media OD was subtracted from the OD for an average 75-year-old observer (values taken from the Pokorny et al., 1987, model for that age). This difference in OD was then subtracted, wavelength by wavelength, from the log10 of the energy spectrum emitted by the CRT phosphors for the simulated chip, which was measured with a spectroradiometer (Spectra-Pritchard PR703A, Photo Research Inc., Chatsworth, CA) for each stimulus from the standard condition. The resulting function represents the light that it would be necessary to present to younger observers on the CRT to produce the same retinal stimulation that the older observer would receive from the standard stimulus. The chromaticity of this spectral distribution was calculated, and the proportions of the three phosphor types necessary to produce this chromaticity were computed to yield the stimulus used in the simulated-aging condition. The stimuli in this condition were analogous to the stimuli in Lindsey and Brown’s (2002) simulation, with the exception that we measured the OD of each individual’s ocular media, whereas Lindsey and Brown used standard values. The simulated-youthening stimuli were prepared in a complementary fashion. Standard stimuli were adjusted for older observers, on the basis of their measured ocular media OD, to re-create the retinal stimulation of an observer with the ocular media OD of the average 25-year-old from the Pokorny et al. model.

The color-naming stimuli were presented in circular (4.4° of visual angle) test patches surrounded by a uniform gray field (28.7° × 21.5°) on a Sony G-200 CRT. The chromaticity of the surround was (0.310, 0.316) in CIE 1931 chromaticity coordinates, equivalent to CIE illuminant C. In the standard condition, the luminance of the background was 10 cd/m², and test-patch luminance was set to twice that of the surround. In
the simulated-aging and -youthening conditions, the actual luminance values of the test patch and surround depended on the measured values of ocular media OD. Stimuli were presented in a darkened laboratory at a viewing distance of 63.5 cm. All observers were properly refracted for the test distance using trial lenses. Stimuli were presented monocularly to the eye that was measured for ocular media OD. Each observer participated in two conditions: the standard condition and a simulated-aging or -youthening condition. In each condition, 40 stimuli were presented four times over the course of two sessions. Each observer began with one practice session.

Stimulus presentation was preceded by 10 min of dark adaptation, followed by 3 min of light adaptation to the mean background color for the condition to be tested. On each trial, the simulated Munsell chip was presented for 1 s. Following each stimulus presentation, the observer used a computer mouse to choose a single color name from a provided list of the 11 English BCTs (red, orange, yellow, green, blue, purple, white, black, brown, gray, and pink).

**RESULTS**

Figure 2 displays the color names chosen most often for each of the 40 simulated Munsell chips for younger and older observers in the standard and simulated-aging and -youthening conditions. The results of the simulated-aging condition replicated Lindsey and Brown’s (2002) major experimental findings. Observers in the this condition used the term “green” for two chips in the blue-green category of the Munsell designation that were named “blue” most often in the standard condition. In addition, chips consistently named “purple” in the standard condition were named “pink” more frequently in the simulated-aging condition. The results for the simulated-youthening condition are also consistent with the results and analysis of Lindsey and Brown. Under these conditions, compared with the standard stimulus condition for older observers, there was a shift toward using “blue” more often when naming chips in the blue-green Munsell designation. Older observers in this condition gave the name “yellow” to one chip that was called “orange” in the standard condition. There was also a tendency for these observers to use “purple” for some chips in the simulation that were named “pink” in the standard.

Although the results from our aging and youthening simulations replicated Lindsey and Brown’s (2002) major findings, the comparison of the results from the older and younger observers in the standard condition contradicted the lens-brunescence hypothesis. According to this hypothesis, the older observers’ color-naming responses under standard conditions should have been more similar to the younger observers’ responses in the simulated-aging condition than to the younger observers’ responses in the standard condition. Most important, in the standard condition, older observers should have used “green” for some stimuli designated “blue” by younger observers, just as was seen in simulated aging. However, when older and younger observers were presented physically identical stimuli, color-naming responses were very similar.

Figure 3 shows the proportion of “green” and “blue” responses for the green-yellow, green, blue-green, blue, and purple-blue Munsell colors. The upper panel compares older and younger observers’ responses in the standard stimulus condition. Clearly, there is little difference between the two groups’ responses for these stimuli, especially in the blue-green region. For the stimuli represented in Figure 3, the response proportions for older and younger observers in the standard condition are highly correlated for both “green” ($R^2 = .95$) and “blue” ($R^2 = .99$). The lower panel of Figure 3 compares the results of the older observers in the standard condition and younger observers in the simulated-aging condition. Younger subjects in the simulated-aging condition used “green” most often for several stimuli that older subjects called “blue.” Correlations for color-name proportions for younger observers in the simulated-aging condition and older observers in the standard condition were lower than the correlations between the two groups when they received the same physical stimuli ($R^2$ for “green” = .80 and $R^2$ for “blue” = .81).
The lens-brunescence hypothesis predicts a strong negative relation between use of the term “blue” and ocular media OD for the stimuli tested in this experiment. This prediction is seemingly confirmed by a comparison of younger subjects’ data in the standard and simulated-aging conditions. For younger subjects, the term “blue,” as a proportion of total responses, was not significantly different for older and younger subjects in the standard condition ($p = .49$). Figure 4 shows the proportion of all trials in the standard condition in which each observer selected the color term “blue” to describe the stimulus as a function of that observer’s ocular media OD at 400 nm. The regression was not significant. The slope of the regression line was $-0.001$ ($R^2 = .009$, $p = .89$). We cannot reject the null hypothesis that there is no relation between ocular media OD and use of the color term “blue” in normal aging.

**DISCUSSION**

Lindsey and Brown (2002) highlighted the interesting correlation between high UV-B exposure and the absence of a verbal

![Figure 3](image-url)
distinction between the color terms “blue” and “green.” This association is particularly interesting as high UV-B exposure is also correlated with accelerated lens brunescence (Javitt & Taylor, 1994; Werner et al., 1990; Young, 1991). Lindsey and Brown suggested a potential physiological mechanism linking these two correlations. They hypothesized that, because short-wavelength light is absorbed by the crystalline lens in greater proportion with increasing UV-B exposure, people chronically exposed to high levels of UV-B light do not require the verbal distinction between blue and green. This model offers a potentially elegant explanation for the geographic distribution of blue-green relative to grue languages based entirely on a simple physiological process.

Our replication of Lindsey and Brown’s (2002) lens-brunescence simulation confirmed their results. When younger observers were presented filtered stimuli simulating the spectral composition of light that would reach an older observer’s eye for the same stimulus, they used the color term blue less often than in the standard condition and used the color term green in its place. However, when the same physical stimuli were presented to younger and older observers, the two groups named the relevant stimuli in virtually the same way. There was no significant relation between ocular media OD and use of the color terms blue and green. This result fails to support the lens-brunescence hypothesis. We infer from our results that premature lens brunescence due to increased UV-B exposure could not serve as a causal mechanism to explain the lack of a linguistic distinction between blue and green.

The older observers in our study had ocular media absorption values that were, on average, approximately 1 log unit greater at 400 nm than those of the younger observers. The measured difference between the most and least dense ocular media in our study was 1.6 log units. The ocular media of the young observer with the least opaque lens and cornea transmitted 41 times more light at 400 nm than the ocular media of the older observer with the most opaque lens and cornea. Despite these large differences in the filtering of short-wavelength light, younger and older observers used the terms blue and green in much the same way for physically identical stimuli.

People differ widely in ocular media OD, both across the life span and among individuals of similar age. Despite these large individual differences, people within the same culture describe physically identical stimuli similarly on many measures of color appearance, including unique hue loci (Schefrin & Werner, 1990) and hue scaling (Schefrin & Werner, 1993). Additionally, differences in light-source spectral composition across viewing situations (e.g., the sun vs. a fluorescent lightbulb) are such that the light reaching the retina from any given object will differ greatly from situation to situation. Despite the vastly different spectral composition of light reaching the retina across the life span and across different viewing conditions, people generally refer to a given object with a particular color term. If this were not the case, color names would not be useful linguistic tools.

For this uniformity to be possible, the visual system must compensate for changes in the spectral composition of light reaching the retina. This compensatory process is referred to as color constancy and has been the subject of systematic theoretical and experimental investigation for well over 100 years (e.g., Helmholtz, 1911/1962; Hering, 1920/1964). Although the mechanisms underlying this process are not fully understood, chromatic adaptation and surround effects likely play critical roles (Kraft & Brainard, 1999). An individual’s entire visual world is filtered through his or her ocular media. The visual system is likely able to discount the spectral filtering effects of ocular media because of this consistency across space and over time.

Lindsey and Brown (2002) referenced an apparently competing sociolinguistic explanation for the absence of a blue/green lexical distinction in languages of groups living in high-UV-B regions (Berlin & Kay, 1969; Kay & Maffi, 1999). According to this hypothesis, a driving force in increasing the number of color terms over time could be, as Lindsey and Brown acknowledged, that “as a culture becomes technologically more complex, speakers have more frequent need to distinguish objects by their colors” (p. 512). This argument is supported by the observation that distance from the equator, level of technology, and number of BCTs are all positively correlated (Ember, 1978; Hays, Margolis, Naroll, & Perkins, 1972; Naroll, 1970). However, the reason tropical, and hence high-UV-B, societies tend to be less technologically advanced than other societies and vice versa, although unexplained, is beside the point currently at issue.

Despite Lindsey and Brown’s (2002) welcome attempt to offer a physiological mechanism for the lack of a lexical blue/green distinction, other aspects of observed cross-language color naming, beyond the experimental data presented here, pose a problem for their hypothesis. The most notable characteristic of
the color-term systems of the languages of low-technology groups is not so much the lexical merger of green and blue per se as it is the merger of several specific pairs and triples of basic colors into single terms (Kay & Maffi, 1999). For example, although the inclusion of red and yellow in a single term is less frequent than the inclusion of green and blue in a single term, languages that contain these two mergers are similarly distributed across the globe (see Lindsey & Brown, 2002, Fig. 2). The merger of red and yellow could not be explained by increased lens brunescent, because the wavelengths critical for this distinction pass through the ocular media virtually unfiltered. A model of color-name evolution must account not only for the geographic distribution of languages that lack distinct terms for blue and green, but also for the distribution of languages that lack other lexical distinctions, notably, the red/yellow distinction. The correlation between UV-B radiation and the absence of terms meaning blue is likely not caused by accelerated brunescent of the ocular lens, but is likely related to some additional third factor that is correlated with both phenomena.

Acknowledgments—This work was supported by the National Institute on Aging (Grant AG04058), the National Science Foundation (Grant 0130420), and a Jules and Doris Stein Research to Prevent Blindness Professorship.

REFERENCES


(Received 3/8/04; Accepted 4/19/04)
This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.