

Variations in daylight as a contextual cue for estimating season, time of day, and weather conditions

Jeroen J. M. Granzier

Justus Liebig University, Department of Psychology,
Giessen, Germany



Matteo Valsecchi

Justus Liebig University, Department of Psychology,
Giessen, Germany



Experience and experiments on human color constancy (i.e., Arend & Reeves, 1986; Craven & Foster, 1992) tell us that we are capable of judging the illumination. However, when asked to make a match of the illuminant's color and brightness, human observers seem to be quite poor (Granzier, Brenner, & Smeets, 2009a). Here we investigate whether human observers use (rather than match) daylight for estimating ecologically important dimensions: time of year, time of day, and outdoor temperature. In the first three experiments we had our observers evaluate calibrated color images of an outdoor urban scene acquired throughout a year. Although some observers could estimate the month and the temperature, overall they were quite poor at judging the time of day. In particular, observers were not able to discriminate between morning and afternoon pictures even when they were allowed to compare multiple images captured on the same day (Experiment 3). However, observers could distinguish between midday and sunset and sunrise daylight. Classification analysis showed that, given a perfect knowledge of its variation, an ideal observer could have performed the task over chance only considering the average chromatic variation in the picture. Instead, our observers reported using shadows to detect the position of the sun in order to estimate the time of day. However, this information is highly unreliable without knowledge of the orientation of the scene. In Experiment 4 we used an LED chamber in order to present our observers with lights whose chromaticity and illuminance varied along the daylight locus, thus isolating the light cues from the sun position cue. We conclude that discriminating the slight variations in chromaticity and brightness, which potentially distinguish morning and afternoon illuminations, lies beyond the ability of human observers.

Introduction

Looking at our own phenomenological experience we have a feeling that we are capable of perceiving changes in daylight. For example, at sunset, we sense a reddish glow to things. By looking at the same outdoor scenes at various times of the day (see Figure 1), through seasonal changes and under overcast or sunny skies, one realizes that the same object can appear to be of a different color under different natural illuminants. An object's changes in color under shifts in natural daylight can, however, be very sophisticated to the untrained eye. Painters such as Monet used this fact to great effect in their paintings. The chromaticity, brightness, color contrast, and brightness contrast of surfaces in a scene might help the observer in making inferences about the nature of the illuminant (Brenner & Nascimento, 2012; see also Figures 2 and 3). These changes in daylight illumination are best observed through the window of a darkened room, as Monet did when he painted his series of Rouen Cathedral facades. The color changes range clearly from warm (yellow) to cool (blue). As the sun declines in the sky, the light dims and the sky color shifts from deep blue to cerulean; in surface colors, reds and yellows become more saturated, yellow greens become warmer and lighter valued, and blues or blue greens become grayer and darker. These changes physically occur because the daylight spectral power distribution contains different proportions of long wavelength (red), middle-wavelength (yellow), and short wavelengths (blue) at different times of the day, different seasons of the year, and different geographical locations, as well as under different atmospheric conditions.

Although the color of objects might change slightly depending on the change in (daylight) illumination, we

Citation: Granzier, J. J. M., & Valsecchi, M. (2014). Variations in daylight as a contextual cue for estimating season, time of day, and weather conditions. *Journal of Vision*, 14(1):22, 1–23. <http://www.journalofvision.org/content/14/1/22>, doi: 10.1167/14.1.22.



Figure 1. Examples of images showing that we are able to perceive changes in daylight. Photos shown were shot at the first author's garden taken on the same day in May but at different times during the day, ranging from 6 a.m. to 9 p.m. Photos by Marjon Driessen.

are quite good in attributing these differences in objects color appearance to changes in the illumination, a phenomenon called color constancy (Arend & Reeves, 1986; Craven & Foster, 1992). We learn to judge how an object would look under different illuminations, and since our interest lies mainly in the object's color, we become unconscious of the sensations on which these color constancy judgments rests.

Experiments, originally performed by Arend and Reeves (1986), indicate that illuminant differences between otherwise identical scenes (computer-simulated two-dimensional "Mondrians") are readily visible even without reference surfaces available that provide conclusive cues. Furthermore, their results indicate that (some) observers can use cues to the illuminant color

from reference surfaces to improve color constancy, while others find this task difficult to perform (Cornelissen & Brenner, 1995; Granzier, Vergne, & Gegenfurtner, 2013). These results suggest that illumination perception and color constancy are strongly linked. However, when it was explicitly tested to determine whether such a link exists, the results seem to indicate otherwise (Granzier, Brenner, & Smeets, 2009a; Granzier, Nijboer, Smeets, & Brenner, 2005; Granzier, Smeets, & Brenner, 2012). These results make the precise relationship between color constancy and illumination perception complex. The results of Arend and Reeves (1986) introduced the interesting issue whether an observer can represent, simultaneously, the color of a surface and that of the light illuminating it



A: Sunrise.



B: Noon.



C: Sunset.

Figure 2. Images illustrating the various colors and intensities of illumination at sunrise (A), noon (B), and sunset (C) during the same day. Notice the significant change in the colors of the sky. Please note that the colors of the images in print might appear to be different from the calibrated images as shown during the experiment.

(Arend, 1994; MacLeod, 2003; Mausfeld 1998). There is to date little experimental evidence for such multidimensional perceptual responses for chromatic scenes. The focus of the current investigation, however, is on daylight brightness and chromaticity variations and whether observers are able to use the brightness and chromaticity information to predict time of day and season or month. Thus, the term “illumination” in this paper only refers to its chromaticity and brightness (not distribution, flow, etc.) and we do not address, for instance, the angular distribution of the luminance.

Two different and opposing theories about the role of illuminant perception in human color perception have been postulated. On one hand, there are models that assume that illuminant estimations are based on an unconscious, automatic process that is achieved in the earliest stages of visual processing (Foster & Nascimento, 1994; Land & McCann, 1971; Von Kries, 1905). Estimations about illumination are considered to be

only useful for achieving a stable objects’ color representation under changes in illumination. The contribution of the illumination in the light signal (i.e., the light reaching the eyes) are estimated by the visual system and then discounted. The implicit assumption in these models is to regard deviations from perfect color constancy as mere mechanical limitations of the visual system that we should interpret as errors. We will refer to this hypothesis as the *illuminant estimation hypothesis* (Beck, 1972; Epstein, 1973; Koffka, 1935). Many computational theories of color constancy are based on this hypothesis (i.e., Brainard & Freeman, 1997; Buchsbaum, 1980; D’Zmura & Lennie, 1986). Thus, an implicit assumption in these models is that we are unaware or unconscious of the illuminants’ chromaticity and brightness.

On the other hand, having a conscious representation of the illumination itself may be of interest. As stated above, our phenomenological experiences tell us



A



B

Figure 3. Images taken at the same time of day, but on days in summer (A) and fall (B). Images were taken on clear and sunny days. Notice the subtle differences in colors and the positions of shadows. Please note that the colors of the images in print might appear to be different from the calibrated images as shown during the experiment.

that the colors of objects *do* change slightly under changes in illumination. This might indicate that we do have a perception of the illuminant. However, these slight shifts in an object's color appearance should not matter much for identifying an object based on its color, as long as the visual system is able to attribute these shifts to the changes in illumination (Arend & Reeves, 1986; Craven & Foster, 1992). This raises the fundamental question of whether perceiving the illumination in itself would be advantageous for an individual. Rather than discounting the illuminant, would it not be more desirable to recognize that surfaces were being viewed under different illuminants, to infer the relative properties of different illuminants, and to identify surfaces across illuminant changes (e.g., Zaidi, 2001)? We set out to test one possible use of perceptual information about the changes and statistical regularities of light, separate from the information

about the relatively constant physical properties of objects within a scene. In ecological terms, information about the illumination may be useful when assessing, among other things, weather conditions (Endler, 1993; Jameson & Hurvich, 1989; Zaidi, 1998) and time of day (Cochran, Mouritsen, & Wikelski, 2004). Thus although the suggestion has been made that perceiving the illumination in itself might be useful, to date no empirical studies have been carried out that have investigated this claim in human observers.

Here it is worth mentioning that studies that have tested the amount of color constancy in human observers found that, although color constancy is on average quite robust (Granzier et al., 2009a, 2009b), it is far from perfect (for an overview see Foster, 2011). One could even speculate whether the reason for this imperfect color constancy might perhaps be to keep some information about the illuminant to guide the organisms' behavior, assuming that lightness and illumination cannot be represented independently. Thus, it might be the case that discussions about illumination perception are only useful (or relevant) in the context of ecologically valid questions (predicting or estimating the weather, time of day, and the season or month). This might explain why asking observers to match the color and brightness of the illumination in a real scene leads to poor performance as this task does not resemble in any way the kind of problem-solving skills for which our illumination perception abilities were “designed.” Here we test illumination perception by using a different approach and try to answer the question whether illumination perception is used by the visual system to answer ecologically valid questions, such as estimating the month, time of day, and weather conditions. For correctly estimating these features of a scene, a correct representation of the illumination is needed.

Up to this point we have been discussing the concept of color constancy as a single phenomenon. However, color constancy might be like a “bag of tricks”; the kind of information and the combination of information that will be used by the visual system will depend on the task at hand (i.e., Brenner, Granzier, & Smeets, 2011; Granzier, Smeets, & Brenner, 2006), the observer (Granzier & Gegenfurtner, 2012; Granzier, Toscani, & Gegenfurtner, 2012), and the presence of the information itself (for an overview see Foster, 2011; Smithson, 2005). Looked at it in this way, color constancy is like intelligence: we all know what we mean if we talk about it, but it is hard to define precisely what it actually is. Indeed, as already stated above, it might even be the case that the visual system does not need to make an estimate of the illumination in order to achieve color constancy. Cone adaptation could, to some extent, already lead to a form of color constancy.

Variations in daylight

The spectral power distribution of light emitted by the sun is almost invariant, and the range of daylight we experience results from sunlight interacting with the Earth's atmosphere (Henderson, 1977). The spectral distribution of daylight varies considerably depending upon latitude, time of day, season of the year, and weather conditions (Judd, MacAdam, & Wyszecki, 1964; Lee & Hernández-Andrés, 2005a, 2005b). On a sunny day, it consists in the shade of a mixture of the bluish scattered light from the sky and (inter)reflections from the environment. In the open the spectral distribution of daylight consists of a mixture of skylight with the direct rays of the sun. Near sunset, as we all (probably) know, daylight can become very reddish as the sun's slanting rays must pass through an increasing thickness of atmosphere. Changes in the sun's declination causes seasonal changes in the sun's maximum elevation above the horizon, causing the average illumination and temperature to increase from winter to summer. The atmosphere produces changes in illumination intensity and color through the filtering effects of smoke, dust, water vapors, and clouds.

Blackbody color

It would be useful to find a method that can describe or define the relative amounts of yellow or blue bias in a "white" light. A simple way to do this, for natural light is by the light's blackbody temperature. In 1900 the Austrian physicist Max Planck mathematically described the spectral power distribution that would be produced at different temperatures by a perfectly radiating object, called a "black body" because no light would reflect from it. These blackbody curves approximately match the spectra radiated by many natural light sources. In all these cases, an entire spectral emission curve can be specified by its blackbody temperature alone.

Correlated color temperature

The blackbody locus provides the method necessary to specify the color of almost any naturally occurring light source. The temperature (curve shape) of the blackbody is adjusted until its standardized spectral emission curve produces a *visual* or metameric match between the blackbody and light source; their chromaticity points are the same. Then the temperature of the blackbody curve, expressed in degrees Kelvin (K), is the correlated color temperature (CCT) of the matching light.

Solar and daylight color

How well do correlated color temperatures describe the chromaticities of actual landscape illumination? The blackbody locus closely parallels the aggregate chromaticity variations across a large sample of daylight spectra, measured in different sky directions across different season and geographic regions at different times of day. In general, there is an extremely close fit between the daylight and blackbody curves. This is not surprising, because the solar spectrum is one of many natural light sources that resemble a blackbody radiator. But the most essential point is that the blackbody locus describes the entire sequence of landscape illumination across diurnal and seasonal cycles.

The chromatic changes in daylight correspond to correlated color temperature changes within the approximate range of 40000 K–4000 K (Hernandez-Andres, Romero, Nieves, & Lee, 2001; Judd et al., 1964; Lee, 1994). Some investigators have studied daylight over many years and on the basis of their records it is possible to establish averages that have been proven to be useful for practical applications (Henderson, 1977; Walsh, 1961).

Many aspects of color vision have been attributed to adaptations to the natural color environment. For example, unique blue and yellow lie very close to the axis along which natural daylights vary and may thus reflect a perceptual representation of the daylight locus (Lee, 1990; Mollon, 2006; Shepard, 1992). Similarly, basic color terms, which are the primary landmarks of how colors are named by a language, have been predicted by analyzing the distribution of color or lightness levels in images (Attewell & Baddeley, 2007; Yendrikhovskij, 2001). Thus, there is the argument that the evolution of all visual sensory mechanisms must have been strongly influenced by the characteristics of terrestrial illumination (Shepard, 1992). Of course, over a 24-hour cycle, there is also an overall change in the intensity of the illumination from the very bright midday to the darkest starlit night. This intensity change is claimed to have provided the need for the extremely wide range of luminances to which the visual system is capable of responding (Shepard, 1992).

Variability in scene appearance

The appearance of a fixed scene depends on several factors including the viewing geometry, illumination geometry and spectrum, scene structure, reflectance (BRDF) and the atmosphere (fog, rain, etc.) in which the scene is immersed. The distribution of daylight illumination on a scene produces a wide variety of scene appearances. Figures 3 and 4 illustrate the various

shadow configurations on a sunny day. Figure 2 shows the different illumination colors and intensities at sunrise, noon, and sunset. The types of outdoor illumination and weather conditions change with season. For instance, the intensity distribution of sunlight and skylight differ from summer to winter (Henderson, 1977). Similarly, the atmospheric conditions that can be seen in fall are significantly different from those that occur in summer (see Figure 2). There have been very few studies about how the colors of landscapes change during a seasonal cycle (Hering, 2007; Webster, Mizokami, & Webster, 2007; Webster & Mollon, 1997). For example, Hering (2007) over the course of a whole year matched the color appearance of landscapes by selecting chips of a modified NCS atlas. Measurements were done in the Heuckenlock Nature reserve, Germany. The author found large variations in the average perceived chromaticity of the same landscapes and surfaces depending on the season. Similar results have been found by Webster et al. (2007). Thus, seasonal climate changes alter both the average color in the scenes and how colors are distributed around the average.

Webster et al. (2007) also found that the long-term changes in environmental color at each measured location are primarily due to changes in the reflectances of the surfaces rather than to changes in the daylight loci, although changes in daylight loci did change the light signal to some extent, depending on the season. However, all measurements were done at the same time of day (midday) and therefore did not reflect the full range of daily variations in illuminant chromaticity. Another problem with the latter measurements is that the spectral measurements of the landscapes were taken with the camera directed away from the sun. This method could therefore also have underestimated the variability in the light signal caused by the variation in daylight loci. We are not aware of any reports in which daylight spectral measurements have been presented as a function of season and time of day (i.e., daytime hour).

The light field

The light field is a function that describes the amount of light faring in every direction through every point in space. Koenderink, Pont, van Doorn, Kappers, & Todd, (2007) provided evidence that observers have a mental representation of what these authors call “the physical light field.” In these experiments they inserted in the center of a stereoscopically presented three-dimensional scene, a white “gauge” sphere that observers could adjust to match the (a) direction of the light, (b) the diffuseness of the light, and (c) the intensity of the light of the scene. By moving the sphere



A



B



C

Figure 4. Images illustrating different shadow configurations on a clear and sunny day at 10 a.m. (A), noon (B), and 2 p.m. (C). Shadows provide cues for illumination direction. Notice the positions of the sharp shadows on the buildings.

around in space, they found that observers were quite sensitive to these various parameters of the physical light field and generally arrived at close to veridical settings. These results suggests that observers have implicit expectations concerning how objects should

appear in three-dimensional scenes, and that these expectations are measurable. Thus, Koenderink and colleagues demonstrated that observers have representations of both light intensities and the direction(s) of the light source(s) throughout space. As far as we can tell, most of the studies that have tested illumination perception have been done in the lightness domain (see Schirillo, 2013 for a recent overview). We therefore lack an understanding of how observers infer the illumination in a chromatic domain. As already stated above, the current investigation is not concerned with testing observers' ability to perceive the light field.

Image databases

For testing the ability of observers to estimate time of day, month, and weather conditions, a large image database is needed that has captured images of outdoor scenes during different times of days and different months. Obviously, a ground truth (the actual day, month, and weather conditions) must be registered. Several databases of images of outdoor scenes have been collected. The “natural stimuli” collection (Van Hateren & van der Schaaf, 1998) has around 4,000 images of natural scenes taken on clear, foggy, and hazy days. The MIT city scanning project (Teller et al., 2001) provides a set of 10,000 calibrated images acquired over a wide range of the MIT campus. These databases, however, do not cover the complete appearance variability (due to all outdoor illumination and weather conditions) in any particular scene. Finally, webcams capture images regularly over long periods of time. However, they are usually low quality, noncalibrated, not tagged with ground truth data, and only focus on activity in the scene. A database that does meet our requirements is the Weather and Illumination Database (WILD).

Methodology

The WILD database

For our experiments, we used the Weather and Illumination Database (Narasimhan, Wang, & Nayar, 2002), which consists of high quality (1520×1008 pixels, 12 bits per pixel) calibrated color (RGB) images of an outdoor scene captured every hour for over one whole year (see www.cs.columbia.edu/CAVE/software/wild). The dataset covers a wide range of daylight illumination conditions and weather conditions. The scene that is shot is an urban scene in uptown Manhattan with buildings, trees, and sky visible. The distances of these buildings range from about 20 meters

to about 5 kilometers. The large distance range facilitates the observation of weather effects on scene appearance (see Figure 3 for the entire field of view). Weather information is automatically collected from the National Weather Service websites every hour. This includes information about sky condition (sunny, cloudy), weather condition (clear, fog, haze, rain), visibility, temperature, and so forth. Other information that is automatically collected are obviously the time of day (hour) and month (see www.cs.columbia.edu/CAVE/software/wild/videos.php for a nice time lapse video).

For further details about the image acquisition and sensor calibration procedures we refer to Narasimhan, Wang, and Nayar (2002).

Experiment 1

Procedure

The stimuli were presented on a calibrated Samsung Sync Master (1100 MB) monitor ($40 \text{ cm} \times 30 \text{ cm}$, 1280×960 pixels, 85 Hz, 8 bits per gun; Samsung, Seoul, South Korea). The nonlinear relationship between voltage output and luminance was linearized by a color look-up table for each primary. To generate the three red-green-blue (RGB) look-up tables, we measured the luminance of each phosphor at various voltage levels using a Graseby Optronics Model 307 radiometer with a Model 265 photometric filter, and a smooth function was used to interpolate between the measured data. The spectrum of each of the three primaries at their maximum intensity was measured with the Photo Research PR-650 spectroradiometer (Photo Research, Inc., Chatsworth, CA). The obtained spectra were then multiplied with the Judd-revised CIE 1931 color-matching functions (Judd, 1951; Wyszecki & Stiles, 1982) to derive CIE xyY coordinates of the monitor phosphors and reconstruct the CIE coordinates of the images as seen by the observers. Please note that since the spectral sensitivity curves of the sensors of the camera taking the original WILD images are not reported, we cannot ensure that the reproduction of the scene chromaticity was veridical, but only that the scene luminance was proportional to the original one. Secondly, we could not match the perspective from where the calibrated camera was standing with respect to the city landscape (WILD images) and the angle between the observer and the calibrated images presented on the computer monitor.

Observers were sitting in front of the CRT in an otherwise dark room. They were dark adapted for about 10 minutes during which time instructions were given. Observers were instructed that they would see many

photographs of the same scene (one at a time), and that each photograph was taken during a different month and/or different time of day. They were told that a series of questions had to be answered for each photograph and that these questions would appear at the bottom of the screen. Once an image of the WILD database was displayed on the computer monitor, observers were asked to indicate which month (December, March, June, or September) the image displayed. Observers were instructed to indicate the month by typing the corresponding number (1 = December, 2 = March, 3 = June, and 4 = September). Observers could see which answer they had typed and could change the answer if they had made a mistake. Once an answer had been given (by pressing the enter key), a response could not be reversed. Secondly, observers had to indicate the hour of the day (ranging from 9:00 a.m. to 4:00 p.m.). They were instructed to type in the number of the hour that they thought the image was taken. For example, if observers thought that the image was taken at 1:00 p.m., they had to press the “1” key. Once again they could see the answer that they had given and could correct their response. Finally, observers were asked to type in the outside temperature (in degrees Celsius). Thus, if they thought the outside temperature was 18 degrees, they had to type in 18. The questions were always asked in this particular order and the questions were always indicated at the bottom of the screen. The order of the trials was randomized both within and between observers.

The experiment consisted of 96 trials (4 months \times 8-hour intervals \times 3 replications). The 96 WILD images were selected from the four months that were used for our study. Each session contained three blocks of 32 images. After the completion of each block, observers were instructed to take a short break but they had to stay in the experimental room. The whole experiment took about one hour for each observer.

Images

In order to present the WILD images on our CRT monitor, we converted the original High Dynamic Range images into 8-bit precision images scaling each RGB channel by a fixed factor, which was applied equally to all pictures. The scaling factor was set to a value as high as possible with the constraint to avoid trimming in the brightest pictures. Notice that we chose to use this fixed scaling strategy, as opposed to, for instance, tone mapping because we wanted to preserve the general luminance level as a potential source of task-relevant information.

Figures 2 through 4 indicate that the WILD images should provide enough information with respect to the daylight changes. The white surfaces of (parts of) the

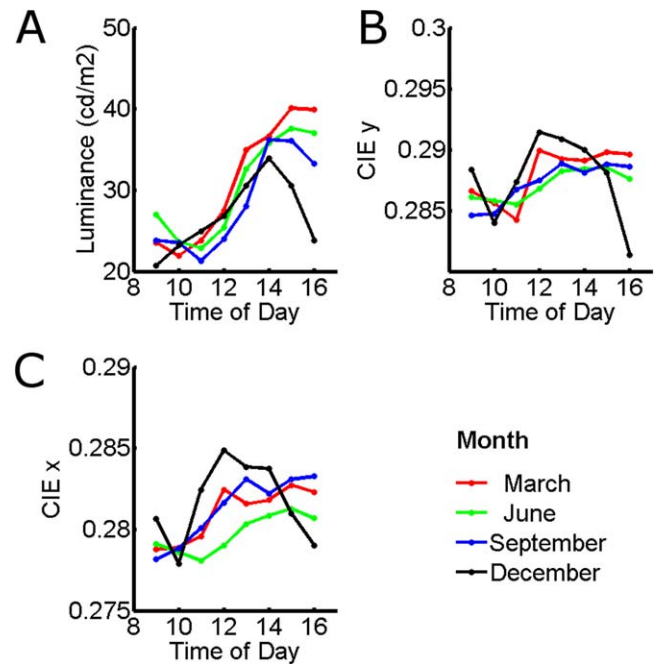


Figure 5. Plotted are the average CIE xy values in the images of the WILD database as they were presented to our observers, separate for each month and time of day (9 a.m. to 4 p.m.). Please note that the general pattern of results appears to be dissimilar between December and the other three months. There might be two reasons for this: the 7-hour interval (9 a.m. to 4 p.m.) covers the daylight excursion in December to a much larger extent, as compared to the other months. Secondly, the changes in chromaticity are minimal and are expanded by the axis scaling.

buildings, the parts of the image that show the sky and many surfaces that are oriented differently so shadows might provide important cues to the position of the sun.

Figure 5 shows the average luminance (cd/m^2) and CIE xy values for the WILD images used in our experiment, shown separately for the images taken during the different time of day and during the different months (shown in a different color). The plots in Figure 5 evidently show that the average chromaticity in the pictures changed as a function of the time of day. In order to quantify the task-relevant information conveyed by the color changes in the images, we computed the average CIE values of each image (see Figure 5). Based on these data, we decoded both the time of day (Figure 6A) and the month of the picture (Figure 6B) using linear classifiers following a leave-one-out procedure. In both cases the classifier performance was significantly better than the chance level (see Figure 6) estimated by randomly permuting the image labels 500 times ($p < 0.001$ in both cases), thus confirming that an ideal observer (i.e., assuming a perfect ability to perceive the average image color and a detailed knowledge of its time-dependent variation within the stimulus set), could perform the task simply based on

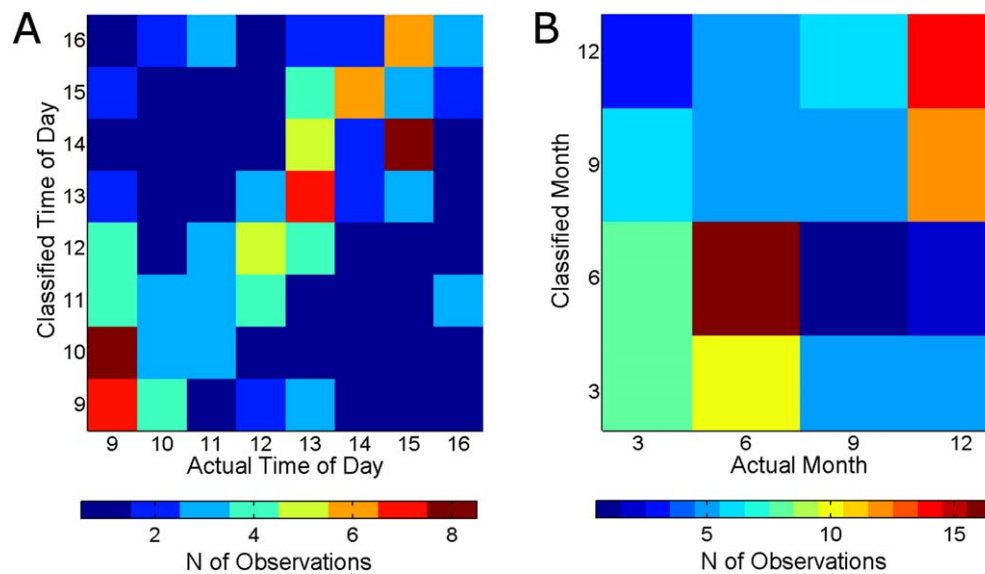


Figure 6. Classification as performed by a linear classifier based on the average CIE values of each image (A: Time of day classification; B: Month classification). In both cases, the classification accuracy was significantly over chance. The total number of images was 96.

these very simple image statistics. In other words, both the classified time of day (Figure 6A) and the classified month (Figure 6B) as shown on the y -axes and the actual time of day and the actual month (as shown on the x -axes) were classified correctly in a significant number of trials, as indicated by the color coding. Although the variability in chromaticity as a function of daylight changes are quite small, they are in the range as reported by others (see for example Webster et al., 2007; Webster & Mollon, 1997). This makes us even more confident that the chromaticity information in the WILD images as displayed during the experiment are physically correctly represented.

Observers

Ten subjects (six males, four females) participated in this experiment. All observers had normal color perception as measured with the Ishihara color plates (Ishihara, 1969). All had normal or corrected-to-normal visual acuity. All observers, with the exception of the first author, were naïve as to the purpose of the experiment. Informed consent was given by all subjects according to the Declaration of Helsinki (World Medical Association, 2004). Methods and procedures were approved by the local ethics committee of the Department of Psychology of the Justus Liebig University.

Analysis

Our main interest was in the number of correct responses. In order to have an indication of how

observers' responses were either correct or biased by the actual month, we plot confusion matrices indicating each observer's choice probabilities as a function of the hour and month the picture was taken. Association strength between the ground truth and the observers' choices was tested with chi-square tests. Furthermore, we binned the pictures based on the average luminance level and investigated the relationship between luminance and both the ground truth hour and the hour reported by each participant.

Results

Time of day

Figure 7 shows the confusion matrix of the observers' hour choices (indicated as "reported time of day") and of the correct hours (depicted as "actual time of day"). The data of each individual are shown in a separate plot. The observer's choices should cluster along the diagonal axis (from bottom-left to top-right) if the two variables were correctly associated. This does not seem to be the case in any observer. In fact, the data of five participants show a significant opposite trend, that is, they tended to indicate that the pictures corresponding to the morning were from the afternoon and vice versa. This result seems to indicate that observers in general used some information within the images to estimate time of day. Obviously, this type of information did not prove to be valid. If observers did not use some information within the images to estimate time of day, the results would have revealed more noise. The results shown in Figure 8 also indicate that there are large interindividual differences in estimates of time of

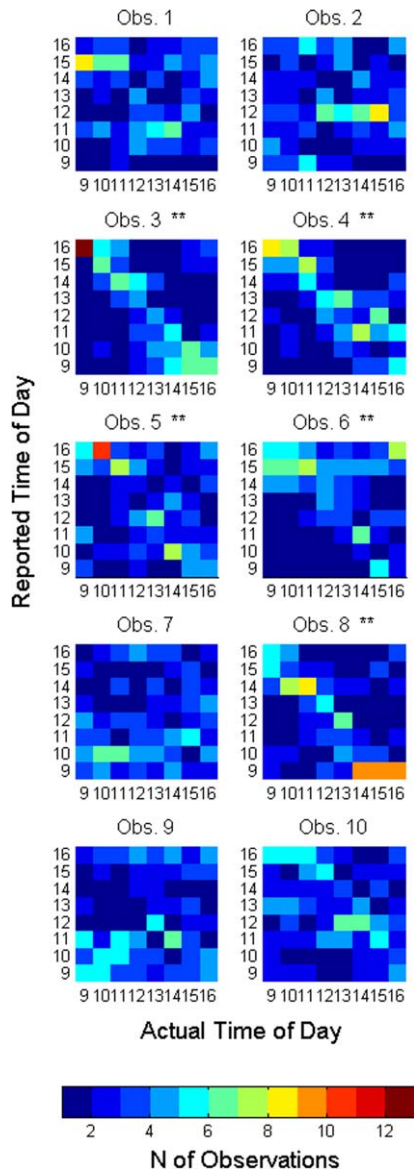


Figure 7. Results of Experiment 1. The color codes represent the absolute frequency with which each observer attributed the pictures to a certain hour (y -axis), depending on the ground truth hour (x -axis). Observers whose data showed a significant association (chi-square test, $p < 0.01$ Bonferroni corrected) are marked by **. If observers would be able to estimate time of day perfectly, the observations would cluster along the diagonal (starting in the bottom-left corner). On the contrary, the observers who showed a significant pattern of association tended to confuse the pictures from the morning and afternoon hours. Total observations, $N = 96$.

day. Thus, for some observers, there does not seem to be any systematic result (i.e., observers 1 and 10), while for others there seems to be the negative correlation mentioned above (i.e., observers 3 and 4).

The average luminance within the image is one source of information that observers might have used,

as discussed in the Introduction, which can indicate the time of day. Sunlight intensity in the pictures we selected from the WILD database was generally lower at sunrise than at sunset, although this was less clear for the December pictures where 4 p.m. is near sunset and luminance decreases again (see Figure 5). In order to investigate whether observers might have used the average luminance in the images to estimate time of day, we plotted the average luminance values of the images as a function of time of day. The results can be shown in Figure 8. Black data points show the actual relationship between the mean luminance of the WILD images and the actual time of day. Red data points indicate the relationship between the mean luminance values of the images and the estimated time of day by the observers. In the data of most participants, we observed a negative correlation between the two variables, and none of the negative correlations was significant. This indicates that most observers have erroneously attributed the brighter images to the morning hours, whereas the brighter images within our subset were on average captured in the afternoon (see Figure 5). Notice that the December data show a weaker correlation (albeit significant) between the ground truth temperature and time of day. Yet, 6 out of the 10 observers showed the negative relationship, including both of the observers where the relationship was significant.

Month

Figure 9 shows the results with respect to estimating the month. These results clearly show large interindividual differences. For example, the data of observer 1 (indicated by “obs. 1” at the top of this figure) shows that he or she was able to correctly estimate the months June and December as the boxes on the diagonal line for these months are colored in red. However, for the months March and September, the boxes are colored blue indicating that this observer almost never guessed these months correctly. A similar trend in the data can be observed for the other subjects (see observers 3, 4, 6, and 10). However, for the other observers, the results are rather mixed. For example, observer 2 shows almost perfect estimates for September, but guessed the other months fairly incorrectly.

The pattern of mistakes is also interesting. Looking Figure 9, it seems that most observers were able to estimate the months June and December correctly. However, most observers seemed to confuse December and March and June with September.

Temperature

Figure 10 shows the results for subjects’ estimates of outside temperature. Shown are the estimated outside

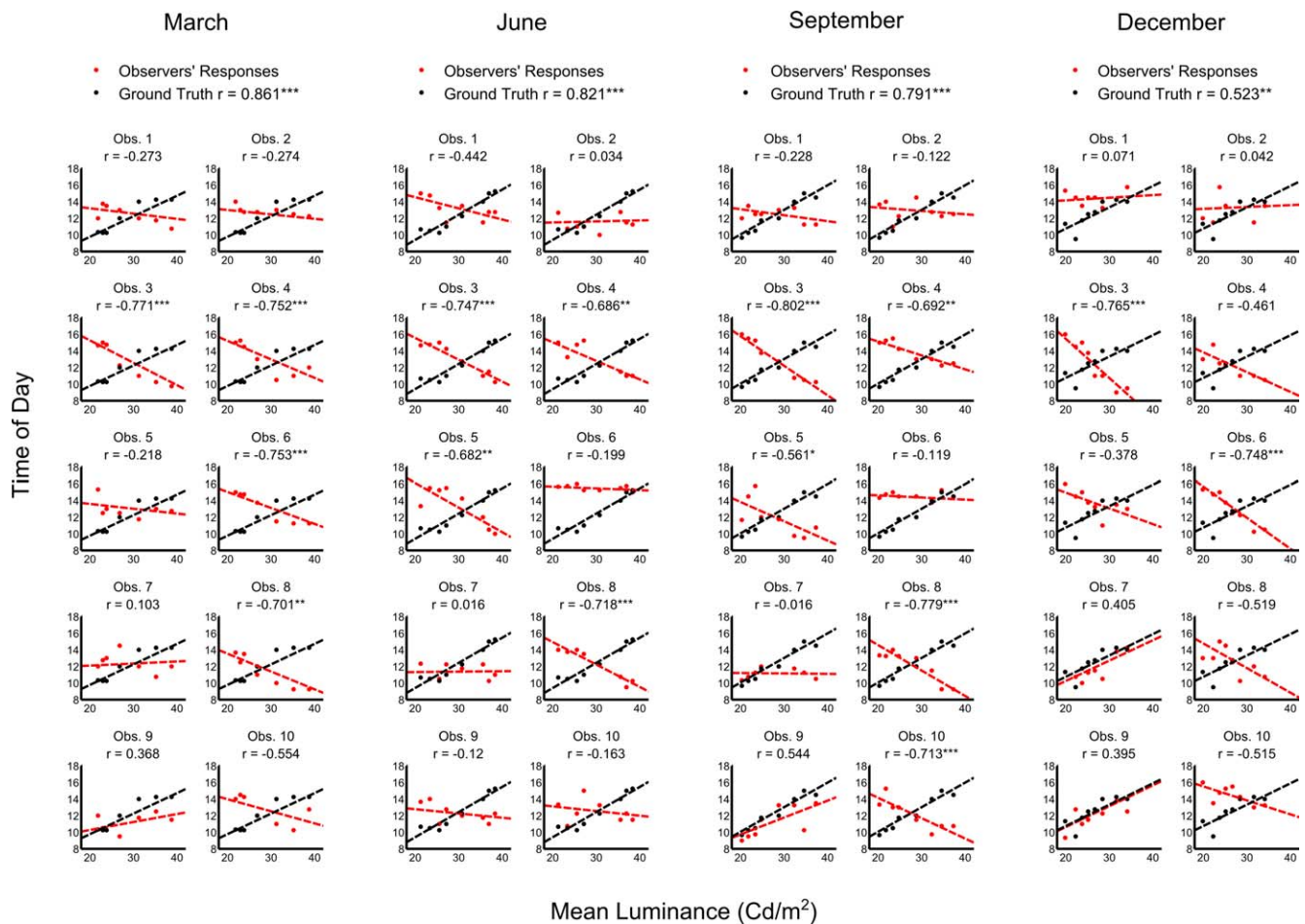


Figure 8. Results of Experiment 1. Reported (red dots) and ground truth (black dots) hours as a function of picture luminance. Whereas the photos with lower luminance were generally captured in the morning hours, a substantial proportion of observers tended to attribute them to the afternoon. The dashed lines are linear fits. Asterisks mark significant correlations (*, **, and *** mark values $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. Single observers' values have been Bonferroni corrected).

temperatures (y-axis) as a function of the actual temperatures (both indicated in degrees Celsius). The data show a macroscopic offset in the temperature estimated by the observers. The estimates are generally above 0°C, whereas the ground truth values were often below the freezing point. This however might be due to the experiment having been conducted in summer or to false assumptions about the location where the WILD pictures were taken. More importantly, at least six observers provided estimates significantly correlated with the ground truth temperature. It is interesting to notice that all of the observers who were able to systematically judge the season (observers 1, 3, 4, 6, and 8) also showed the correct tendency in their temperature judgments. In general, it appears that at least a subset of the observers were able to distinguish the months of December and March (where the average temperatures as the pictures were taken were -3.8°C and -4.6°C, respectively) from the months of June and September (where the temperatures were 12.7°C and

10.7°C, respectively) and estimate the temperature accordingly.

Discussion

It is intriguing that we found a high correlation between the mean luminance of our WILD images and the estimated time of day for some of our observers, with higher luminance indicating an earlier hour of the day. In reality there does not exist such an inverse correlation between luminance and time of day. If anything, mere reasoning alone would indicate that luminance or brightness of a scene would be low both at the beginning and the end of the day reaching a maximum between noon and 2 p.m.; within our stimulus set, luminance increased in general from morning to afternoon. Our observers indicated that the images with the maximum average luminance were captured before noon, which is incorrect. These latter results combined with the pattern of results in general

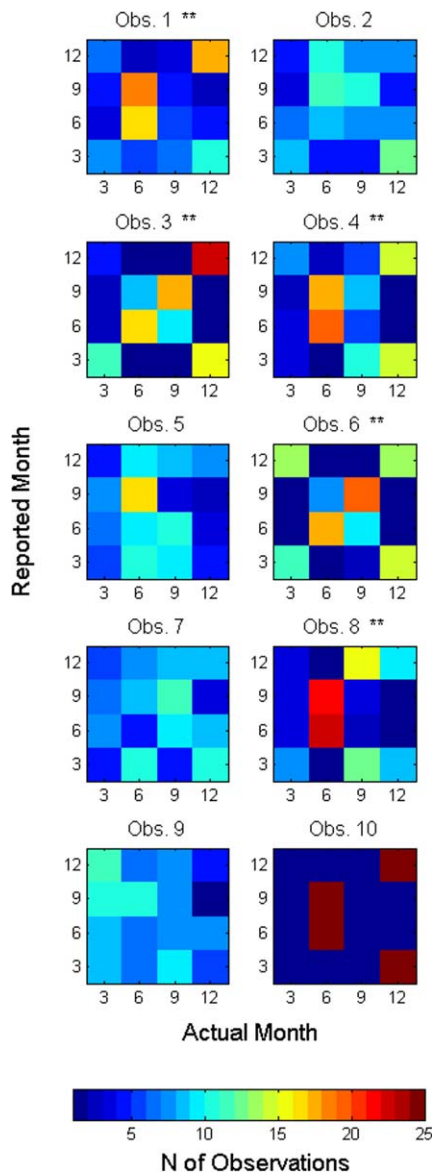


Figure 9. Results of Experiment 1. The rationale of this figure is identical as that of Figure 6B. The frequency with which a month was chosen (shown on the y-axis) is plotted as a function of the actual month (December–March–June–September). Total observations, $N = 96$.

shows that observers possibly used wrong assumptions when judging the time of day.

Overall, we can conclude that observers are also relatively poor in estimating the month although many show over-chance performance. In particular, the months March and September are frequently guessed incorrectly. Some observers seem to have some indication when the months of June and December are displayed. However, there seems to be large between-subject differences in these estimates. Finally, the observers who could estimate the month over chance were also able to estimate the temperature over chance, although in general they were well off the mark.

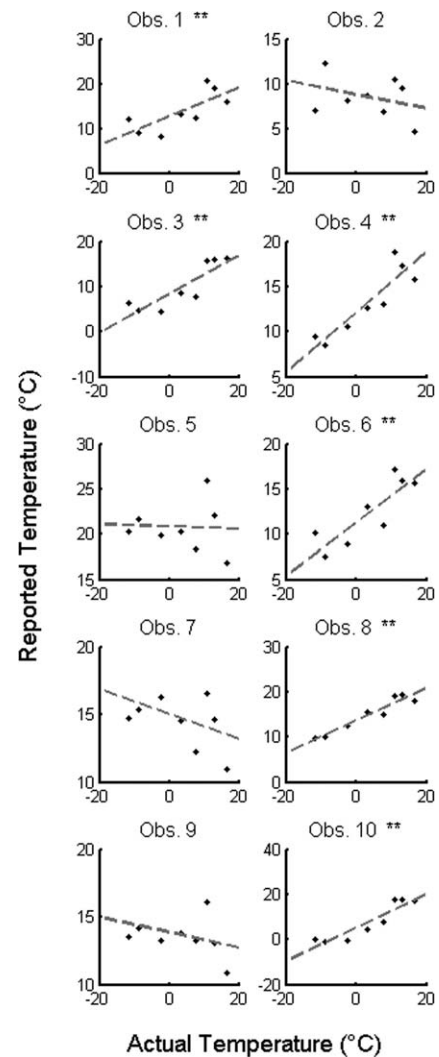


Figure 10. Results of Experiment 1. Estimated outside temperatures as a function of the actual temperatures. The symbol ** denotes a subject whose estimates were correlated with the ground truth. The dashed lines represent the linear regression.

Clearly the time of day estimates of observers were not simply noise, but a systematic but erroneous pattern of results appeared. An important cue for estimating the hour of the day beyond overall brightness, as discussed in the Introduction, are shadows. Shadows give an indication of the positioning of the sun with respect to the Earth’s atmosphere. However, in order for shadows to be a reliable cue for estimating time of day, one has to know where the camera was pointing when the images were taken (north, south, east, or west). It might have been the case that most observers made a wrong assumption with respect to the positioning of the images, which might explain the systematic errors in the observed results. If this assumption is correct, we should find clear deviations in our results if we would mirror the

original images. This is what we set out to do in Experiment 2. In order to avoid any transfer effects from our previous experiment, we used new subjects for Experiment 2.

Experiment 2

The procedure and analysis of this experiment were identical to those of Experiment 1, with the exception of the mirror-imaged WILD database. Observers were not informed about the mirroring of the images.

Observers

Ten subjects (seven females, three males) participated in this experiment. None of these observers had participated in the previous experiment. All observers were naïve as to the purpose of the experiment and had normal color vision, as tested with the Ishihara color plates.

Results

The results of the estimates of time of day are shown in Figure 11. The data are averaged across subjects. The rationale of this figure is identical to that of Figure 7. Clearly the data show no systematic shifts in our results. The results make it unlikely, therefore, that a general assumption about camera direction was shared between observers.

Figure 12 shows the estimates of the month. Similar results were obtained as were found in Experiment 1 (see Figure 9 for more details). June and December were estimated correctly, while the months December and March on the one hand, and the months June and September on the other, were confused regularly.

Experiment 3

In the previous two experiments, most observers only had a very rough sense of time-of-day and month estimates. This might have been the result of the procedure used; we always showed only one image to our observers per trial. Could it be that if more images are presented to observers per trial, and when observers have the choice to rearrange the images by time of day, that estimates are more veridical? This is the question that we set out to study in the current experiment.

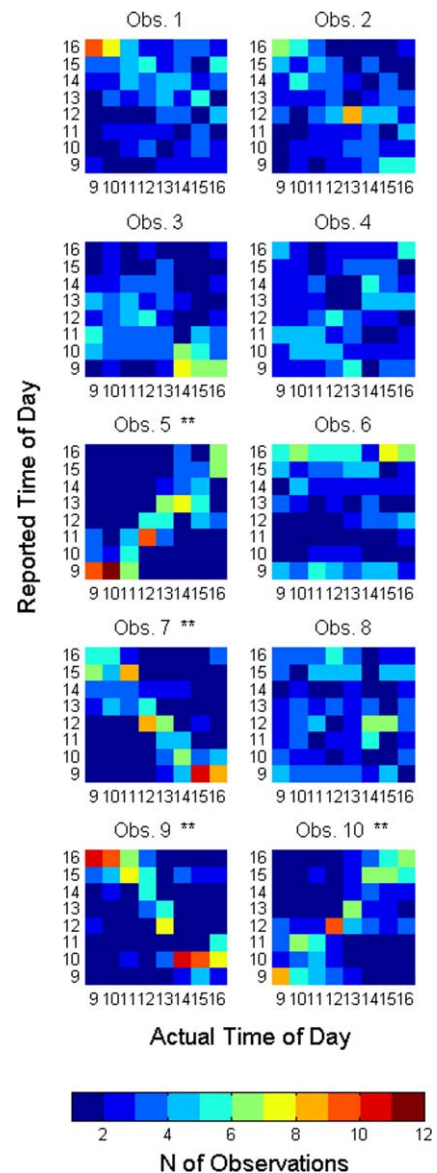


Figure 11. Results of Experiment 2. Confusion matrix of estimated hour of day (shown on the y-axis) and actual hour (plotted on the x-axis). All conventions are as in Figure 7. Using mirror images of the WILD images had a limited effect on the observers' estimates of time of day, only two observers, 5 and 10, produced quite accurate estimates, but observer 7 still interpreted the hour-dependent changes the wrong way and most observers did not produce a clear pattern of results. Total observations, $N = 96$.

Procedure

The same WILD database was used as in Experiment 1. The only difference was that instead of only showing one image per trial, eight images of different times of day (of the same day) were shown simultaneously (ranging from 9:00 a.m. to 4:00 p.m.). Images from the WILD image database were selected from the months March, June, September, and December. For

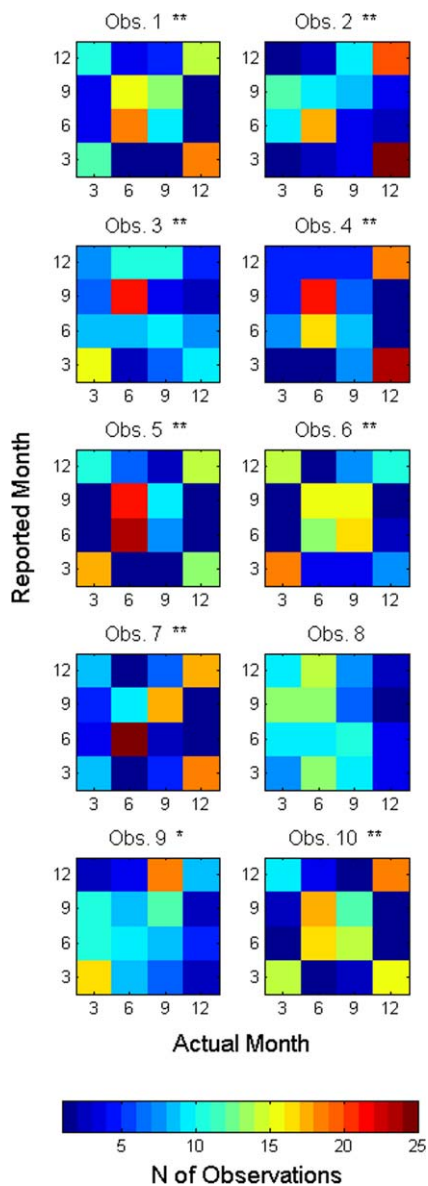


Figure 12. Results of Experiment 2. The estimated month (shown on the y -axis) is plotted as a function of the actual month (plotted on the x -axis), and averaged across observers. For further details see Figure 9. Total observations, $N = 96$.

example, on a given trial, eight images were shown of the different times of day for the third of March. On all given trials, the order in which the images were presented on the screen was randomized. Observers were asked to rearrange the images by arranging the photos from morning to late afternoon (9 a.m., 10 a.m., etc.). They could do this by selecting an image by using the computer mouse and dragging the images in the correct order. Three days were randomly selected for each month, and every trial was repeated twice. This means that for this experiment there were 24 trials in total (4 months \times 3 days \times 2 repetitions). The whole experiment took about 30 minutes.

Observers

Eleven subjects (two males, nine females) participated in this experiment. None of the observers participated in the previous experiments. All observers were naïve as to the purpose of the experiment. All had normal color vision as tested with the Ishihara color plates.

Results

Figure 13 shows the results for Experiment 3. The correct hours of the day are plotted on the x -axis as a function of the estimated hour, which was plotted on the y -axis. The observer's choices should cluster along the diagonal axis from bottom left to top right if the two variables were correctly associated. Even more than in the previous experiments, results show that subjects' estimates of the time of day are not random (as evidenced by a comparison with Figure 7) but significantly cluster around the diagonal lines. Assuming the most likely direction in the relationship within each ordering (i.e., 9 a.m. to 4 p.m. or 4 p.m. to 9 p.m.), 53.2% of the assignments were correct. Randomly assigning the pictures to an hour and computing the most likely direction of the relationship produced a correct performance of 16.3%. All observers except observer 11 showed a level of consistent classification significantly over (95% confidence interval calculated through the aforementioned randomization procedure) the chance level.

For example, observers 1, 4, and 5 show a significant and positive correlation between the actual time of day and the reported time of day. In contrast, several other observers show a highly significant but negative correlation between the actual time of day and the reported time of day (observers 2, 3, 7, 8, and 10). When asked, all observers invariably indicated that they used the position of the shadows in the WILD images to order the pictures in time. These results demonstrate that subjects' inability to estimate time of day in Experiment 1 was probably the result of the observers not having enough information (i.e., only one image per trial) to estimate time of day. Clearly, giving more information with respect to scene appearance at different time of day increases the precision of estimates, although as already stated, many observers are inaccurate, being unable to decide correctly which pictures were taken in the morning and which in the evening.

Experiment 4

Might the insensitivity of our observers with respect to the chromaticity and brightness changes in daylight

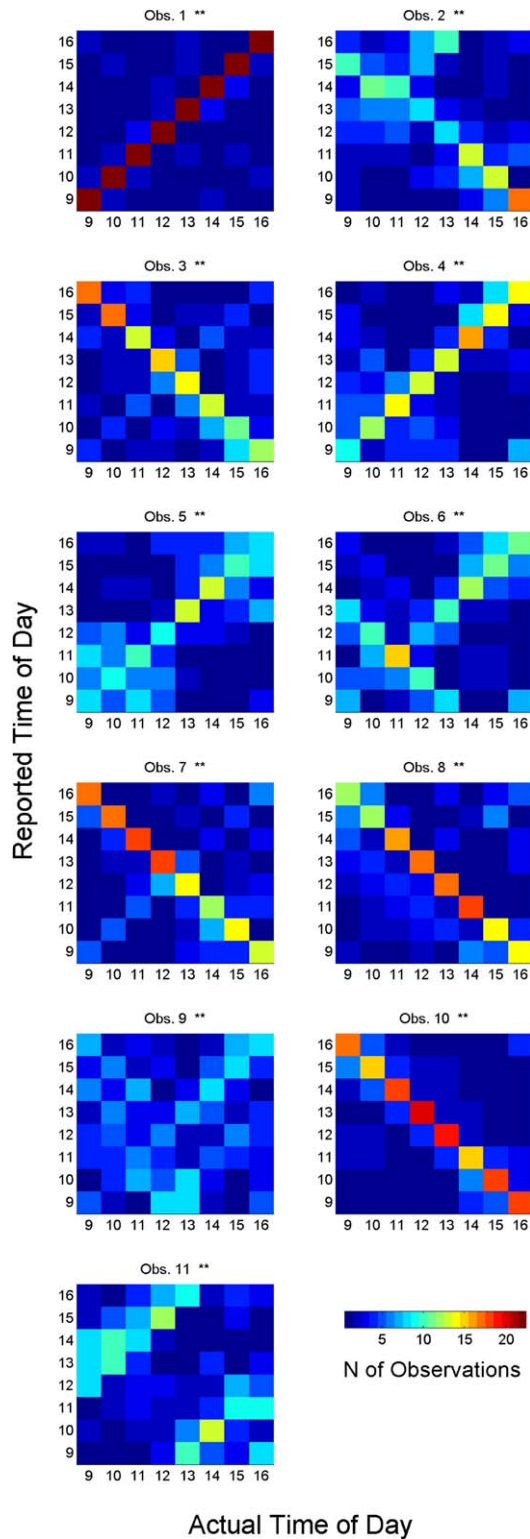


Figure 13. Results of Experiment 3. The actual hour of day (x -axis) is plotted as a function of the reported time of day (y -axis). The data are plotted for each observer separately. For details see Figure 7.

be a result of the fact that observers were pushed to use the position of the sun while looking at the WILD scenes? Moreover, could it be that the variation in the sunlight color and intensity was corrupted by the interaction with the reflecting surfaces in the scenes? As the sun direction changes along the day, different buildings with different albedo are illuminated and this might change the overall chromaticity in ways independent from the illuminant color per se. In order to get rid of these problems we decided to present our observers with daylight illuminations through a LED chamber.

Additionally, is it possible that our observers were unfamiliar with the light variation in different seasons and geographical location? For Experiment 4 we went on to measure the actual variation of sunlight chromaticity and intensity in the same location (Giessen, Germany) and in the same season (June) of testing.

Daylight spectra in Giessen, Germany

As already stated in the Introduction, we are unaware of any reports that have reported the daylight loci as a function of time of day and month. In order to establish whether the range in chromaticity and luminance values that was available in the WILD images (see Figure 5) was representative of natural daylight variations, we measured the daylight loci on several days (in February and June) during different times of day. Results of these measurements can be observed in Figures 14 and 15. The spectrum of daylight reflected by a standard reflector was measured with a PR-650 spectroradiometer (Photo Research, Inc.). The standard reflector was attached to a sundial to obtain daylight measurements when the standard reflector was perpendicular to the position of the sun.

Figures 14 and 15 show that the overall effect of time of day on the daylight loci is quite ambiguous unless one assumes perfect illumination discrimination. For example, if we focus on the luminance information that we have measured (see Figure 14A) one can observe that in the morning and evening the luminance information is almost identical. Thus, observers cannot use this information to determine time of day as this information is ambiguous in itself. A similar explanation can be given for the chromaticity information that is available (see Figure 5). This might explain why observers chose not to use this information even in the restricted case (the WILD images) where the information would have been helpful. When making our daylight loci measurements we had clear sky, which means that the variations in light are solely determined by the sun's angle and by temperature (indeed, the curves in Figure 14 are extremely clean), whereas the

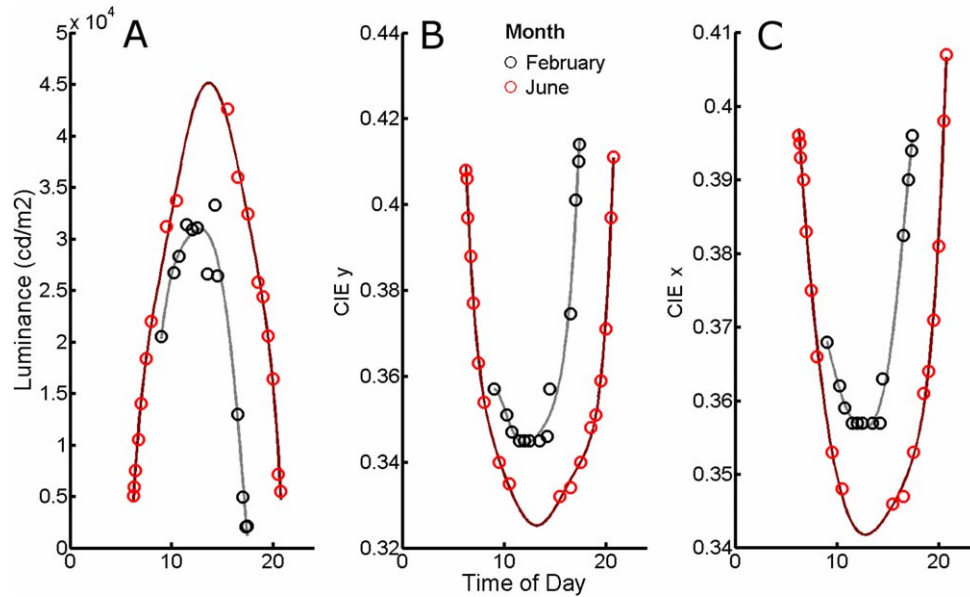


Figure 14. The measured daylight loci are shown plotted in 1931 CIE color space. Measurements were taken on February 20 and 22, 2012 on clear and sunny days. Additional measurements were taken on June 25 and 26, 2012. All measurements were taken at the same location on the campus of Giessen University, Germany, between the hours of 9 a.m. and 7:30 p.m. (February) and between the hours of 6:30 a.m. and 9:00 p.m. (June). Notice the gradual increase in yellow/reddish sunlight at the end of the day (starting from 4 p.m. in summer and at 2:30 p.m. in winter). Measurements were taken by measuring the spectrum of daylight reflected by a standard reflector, measured with a PR-650 spectroradiometer (Photo Research, Inc.). The standard reflector was attached to a sundial to obtain daylight measurements when the standard reflector was perpendicular to the position of the sun. Interpolating lines are sixth degree polynomial fits.

presence of other atmospheric variables (such as clouds) will, if anything, make the information less reliable.

Inspection of the results of Figure 5 and Figure 15 show that there are subtle differences in the chromaticity values as a function of time of day between the WILD images (Figure 5) and the measured chromaticity values of the daylight locus as measured in Giessen (Figure 15). Please note that the WILD images and the measurements that we made in Giessen cannot be directly compared, as the Giessen data were measured using a standard reflector that was always perpendicular with respect to the sunlight. The WILD images, on the other hand, present additional sources of variation in radiance as the sun illuminates the (mostly nonwhite) surfaces of the buildings from a different angle, which in turn can generate a different pattern of interreflections between the surfaces in the scenes. Notice that if one considers only the color of the patch of sky in the top left corner of the WILD images (100×100 pixels), the pattern of chromaticity variation is much more similar (i.e., U-shaped as represented in Figure 5B and 5C) to the one we observed in Giessen (i.e., reddish colors dominate both in the morning and evening hours). On the other side, the luminance excursion in the top left pixels is different from the one we measured in Giessen, since that area of sky

(approximately in the north-west direction) is nearer to the sun toward dusk and luminance increases.

In sum, the daylight loci that we measured at different hours and different months give a straightforward explanation of why our observers could not make reliable estimations of the month and time of day based on changes in daylight, as this information does not seem to be unambiguous. With respect to chromaticity, it is clear that only a perfect knowledge of the daylight color variation, coupled with considerable sensitivity, would allow one to tell morning and evening illuminations apart unambiguously (see Figure 15). The results of our measurements also show that the variation of the daylight locus are largely confined to the hours directly following sunrise and preceding sunset and remain relatively stable throughout the rest of the day, strengthening the suggestion that the effect of illumination on overall image chromaticity could have been largely corrupted by other factors, such as the interplay with colored surfaces in the subset of the WILD database that we used.

The scene

The scene was in front of the subjects, at a distance of 250 cm. The scene had a width of 64 cm and a height of 50 cm. The MacBeth color chart was placed inside

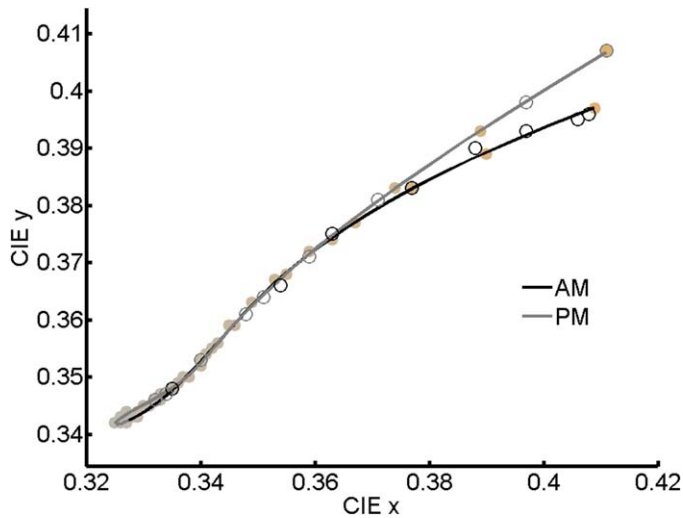


Figure 15. Chromaticity of light in the June Giessen measurements (empty circles) and of the lights used in Experiment 4 (filled dots) in 1931 CIE coordinates. The color of the filled dots gives a rough representation of the color of the light (keeping luminance constant). Evidently the variation of light chromaticity is very similar in the morning and in the evening. However, in principle, an ideal observer with perfect knowledge of the path of the chromaticity variation could tell the early morning illumination apart from the late evening ones. Also, notice that the lights used in Experiment 4 are equally spaced in daytime, indicating that the divergence in chromaticity between morning and evening illuminations is limited to the near temporal vicinity of sunset and sunrise.

the LED chamber that otherwise contained gray surfaces. The MacBeth color chart was developed in 1976 and its 24 patches include six neutral colors, red-green-blue and cyan-magenta-yellow primaries, and other important colors, such as light and dark skin, sky blue, foliage, and so on (McCamy, Marcus, & Davidson, 1976). Please note that the MacBeth color chart includes neutral surfaces providing an abundance of illuminant information. Most of the changes between subsequent illuminants were quite evident to the observers. We did not use familiar objects in the scene for two reasons: First, one study found that the presence of familiar objects does not significantly lead to better matches of the illumination (Granzier et al., 2009a), and second, another study showed that the presence of familiar objects does not significantly help in achieving color constancy (Granzier & Gegenfurtner, 2012). The latter study was still important, as Granzier and colleagues (2009a) showed that there is no correlation between conscious estimates of the illumination and color constancy performance. The presence of familiar objects could therefore still lead to better color constancy performance.

The illumination

We presented our observers with 40 illuminations representing the illumination measured in Giessen in June, at 40 equally spaced points in time between 6:30 a.m. and 9 p.m., as obtained from the sixth degree polynomial interpolation of the observations (see Figures 14 and 15).

Brightness

With respect to the luminance of the LED lamps, we scaled it to fit the gamut of the LED chamber. On average the LED chamber was about 714 times lower (darker) for the chromaticity values used in the experiment compared to the luminance values measured outside at the Giessen University campus. The luminance varied between $7.7 \text{ cd}\cdot\text{m}^{-2}$ and $62.4 \text{ cd}\cdot\text{m}^{-2}$ as measured by taking the luminance of the LED lamps reflected by a standard reflector, and measured with a PR-650 spectroradiometer (Photo Research, Inc.) at the center of the LED chamber.

Chromaticity

The chromaticity of the daylight illuminant was fully covered by the gamut of the LED chamber. The CIE X values were between 0.02 and 0.2. The CIE Y values were in equal steps between coordinates of 0.325 and 0.411. Unfortunately, while the CIE coordinates of the illuminant could be matched, the full spectrum produced by the LEDs in the chamber differs from the one of natural illumination. The result is fully metameric when the illumination is judged based on the chromaticity of the light reflected by the standard white reflector. However, the pattern of colors and luminance values reflected by the Macbeth chart might in principle have given away the fact that the light was artificial. None of the observers complained about the light appearing in any way artificial.

Observers

Eleven observers participated in this experiment (two male, nine female). All were students at the University of Giessen and unaware of the purpose of the experiment. All had normal color vision as tested with the Ishihara color plates.

Procedure

Observers were seated in front of an LED chamber (LED color viewing light; Just NormLicht). The illumination of the LED chamber was under automatic

control by a computer (Dell computer T3500). Subjects were instructed that they would see a LED chamber in which the light represented a daylight illumination that we had measured at the campus in Giessen in the very same month (June). They were told that on each trial they had to indicate the simulated time of day, ranging between 6:30 a.m. and 9 p.m. Furthermore, they were told that the LED chamber would contain colored surfaces displayed on the MacBeth chart (Munsell Color, New Windsor, NY), and that by looking at these samples, they might be better able to indicate the chromaticity of the simulated sunlight. The instructions took about five minutes. Once the instructions were given and subjects were confident that they understood the task, the room lights were turned off and the experimental illumination in the LED chamber was set. A fence was placed at both sides of the LED chamber so that observers could only see the LED chamber. The experimenter and other parts of the room were shielded by the fences so that only the objects placed in the LED chamber could be used to estimate the illumination.

Subjects indicated the time of the day (between 6:30 a.m. and 9 p.m.) that corresponded to the daylight that was shown in the LED chamber. They gave their response verbally. The experimenter confirmed whether this was indeed the time that the subject intended. If this was the case, the experimenter wrote down the number of the hour and the next illumination was set. Please note that we chose to provide subjects the opportunity to see the chromaticity and intensity of the illumination actually changing from trial to trial. We were hoping that this information could help observers in obtaining correct estimates of the daylight changes. The order of the illuminations was randomized. Subjects could take as long as they wanted to make their estimates. No feedback was given as to how accurate observers were in their time estimates. Each session took about 30 minutes for each subject.

Results

The results of two representative observers are shown in Figure 16, which shows the simulated time of day illumination on the x -axis and the reported time of day on the y -axis. If observers were able to use the information with respect to the illumination to estimate time of day, their responses would lie on a diagonal line. Clearly, this is not the case for observer 1 (top panel) or observer 2 (bottom panel). Although the choices between morning and evening look completely random, the observers could identify the lights pertaining to the central hours of the day.

Figure 17 shows the data averaged across all observers. The correct time of day responses are shown on the x -axis. The y -axis represents the reported time of

day responses. Clearly, observers could not distinguish between morning and evening light. Thus, even when using chromaticity and brightness changes along the daylight locus in real three-dimensional scenes, observers are unable to estimate time of day.

General discussion

The results of these experiments show that human observers cannot rely on daylight to estimate the time of day. These results are remarkable, as our own introspection tells us that, to some extent, we do have a representation of the illumination and its change across the day. Figures 1 through 3 are good examples of this. We here show that there are large interindividual differences in the way observers interpret changes in daylight when they have to estimate the time of day and, to a lesser extent, the time of year. It has been suggested that these differences between observers might be explained by the different strategies or knowledge that observers have with respect to object colors and illuminations (i.e., Cornelissen & Brenner, 1995).

The failure to use illumination to estimate the time of day does not imply that observers do not have a representation of daylight illumination at all, as quite a few observers gave time or illumination-related responses, albeit often wrong. Their failure must be the result of the interaction between the observers' assumptions about the usefulness of different cues (i.e., orientations with respect to the sun's position) and the relatively low informativity of sunlight chromaticity and intensity. We would like to point out that the general results and conclusions that can be drawn from both the WILD images and the LED experiments based on the Giessen data are very similar despite the fundamental differences in the experimental paradigms of measuring the daylight locus as a function of time of day and season/month. It is beyond our purpose to test precisely how well observers could perceive the illumination. Experiments that have tested observers' conscious perception of the illumination have failed to find a veridical percept (Granzier et al., 2009a).

An important issue might be the question of how scene content might influence our data. Indeed, the WILD images primarily contain vertical surfaces with deterministic structures and a limited range of materials. How well can illumination be determined from such a scene? Perhaps estimating the illumination would be much better for richer scenes containing three-dimensional objects and objects containing different materials. One way in which we wanted to answer this question was the use of the LED chamber. For one, the LED chamber contained a larger gamut,

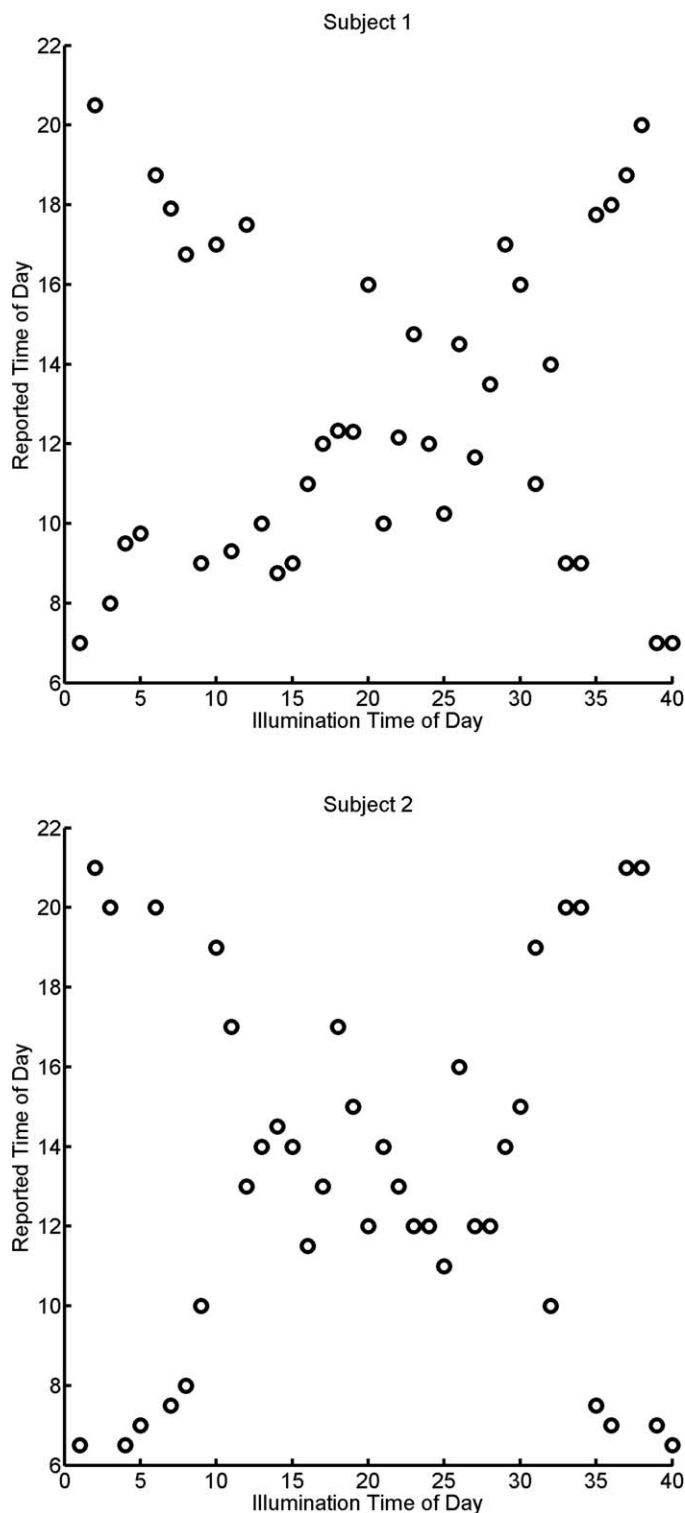


Figure 16. Responses provided by two representative observers in Experiment 4 (LED illumination perception). The observers were able to categorize correctly the central hours of day but failed when they had to distinguish between morning and evening illuminations.

both in chromaticity and in brightness, of the surfaces (the MacBeth color chart).

Another question that might be relevant for our current discussion is whether our results can be extrapolated to conditions in which observers are actually in the light field to be judged. Regrettably, we did not have the technical means to build an illumination chamber that can simulate changes along the daylight locus. However, we would predict that our results would be very similar to those obtained by our current experiments as observers were seated close to the LED chamber in our current experiments (distance of 2.5 m), which guaranteed that the visual field of the perceived illumination was quite large. However, daylight characteristics as a function of location and time of day vary, not only with respect to chromaticity and brightness of the illumination, but also to the average direction of the illumination, the distribution, glare, and other factors with respect to the illumination (see Introduction). Therefore, the visual conditions when the observer is in the light field itself are obviously different compared to when looking at a LED chamber.

Studies comparing color constancy across diverse illuminant changes have drawn an inconclusive picture. Brainard (1998) used two illuminants close to and an additional nine illuminants off the blackbody locus and concluded from his results that the visual system compensates equally well for illumination changes on and off the blackbody locus. However, Ruttiger, Mayser, Serey, and Sharpe (2001) found actually higher color constancy for red–green illuminant changes than for daylight changes. Delahunt and Brainard (2004b) could not report a clear advantage of daylight illuminant changes over other illuminant changes. Daugirdiene, Murray, Vaitkevicius, and Kulikowski (2006) also compared color constancy levels for on- and off-blackbody locus illuminants and did not find superior constancy for the on-blackbody locus illuminants, in line with Hedrich, Bloj, and Ruppertsberg (2009). However, a recent report (Crichton, Pearce, Mackiewicz, Finlayson, & Hurlbert, 2012) measured color constancy for a scene with real objects under a broad range of illuminations, both on and off the daylight locus. They found significantly better color constancy for test illumination chromaticities on the daylight locus. It might be that differences in task procedure and testing a much larger range of illuminations might have caused the differences in results of the latter report. To summarize, it is still unclear whether the visual system can estimate the typical daylight changes more effectively compared to atypical illuminant changes. And, as already stated in the Introduction, the different tasks used to measure color constancy might not only lead to different results, but

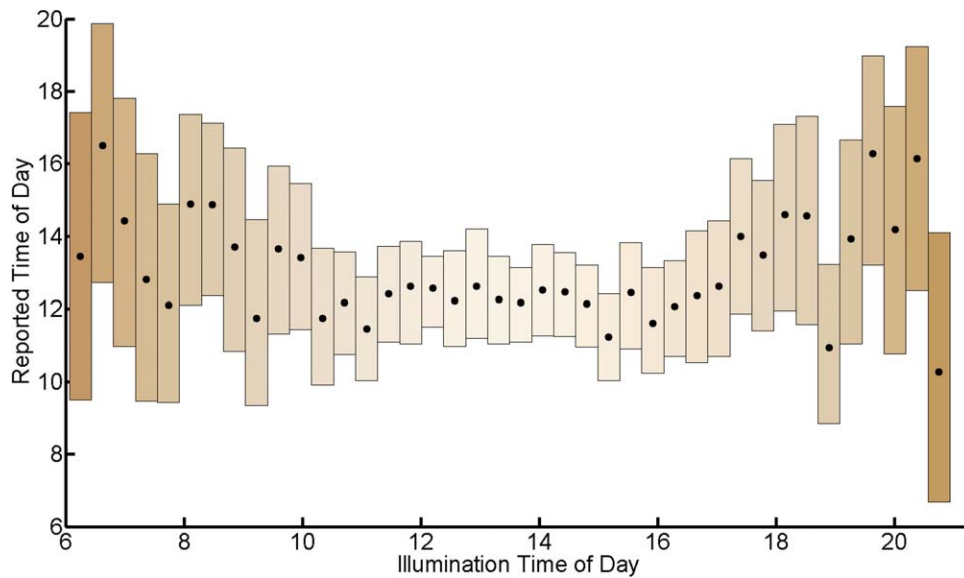


Figure 17. Results of Experiment 4 (LED illumination perception). Black dots indicate the average hour reported by the observers (ranging between 6:30 a.m. and 9 p.m.) as a function of the daylight hour. The vertical rectangles indicate the 95% confidence intervals of the mean across observers, and their color approximates the chromaticity and intensity of the corresponding light.

might even measure completely different phenomena as color constancy is not a unified phenomenon.

Our study is new in that we tested possible ecological reasons for an incomplete discounting of the illumination. Rather than a malfunctioning of the visual system, this could be a potentially important skill if illumination could be used to infer relevant aspects of the environment, such as the time of day. This does not seem to be the case and a possible explanation lies in the information conveyed by daylight. To our knowledge, we measured the changes in chromaticity and intensity of daylight as a function of hour of the day and month (February and June) for the first time. Our measurements indicate that the differences in the chromaticity of the sunlight are very subtle between morning and evening, and observers do not seem to be able to use the illumination for anything other than identifying the central hours of day.

Keywords: illumination perception, daylight, color constancy, color perception, brightness

Acknowledgments

We thank the people from the CAVE Lab of the Computer Science Department of Columbia University that provided us with the WILD images. We also thank Dr. Romain Vergne for providing the software to control the LED chamber and Karl Gegenfurtner for giving us the opportunity to explore our own interests. This research was supported by DGF grant GE 879/9 to KRG.

Commercial relationships: none.

Corresponding author: Jeroen J. M. Granzier.

Email: j_granzier@hotmail.com.

Address: Department of Psychology, Justus Liebig University, Giessen, Germany.

References

- Arend, L. (1994). Surface colors, illumination and surface geometry: Intrinsic image models of human vision. In A. L. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 159–214). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Arend, L. E., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *3*, 1743–1751.
- Attewell, D., & Baddeley, R. (2007). The distribution of reflectances within the visual environment. *Vision Research*, *47*, 548–554.
- Beck, J. (1972). *Surface colour perception*. Ithaca, NY: Cornell University Press.
- Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *15*, 307–325.
- Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *14*(7), 1393–1411.

- Brenner, E., Granzier, J. J., & Smeets, J. B. J. (2011). Color naming reveals our ability to distinguish between a colored background and colored light. *Journal of Vision*, 11(7):8, 1–16, <http://www.journalofvision.org/content/11/7/8>, doi:10.1167/11.7.8. [PubMed] [Article]
- Brenner, E., & Nascimento, S. (2012). Judgments about the intensity of the illumination are influenced by the association between colour and luminance in the scene. In *Proceedings of the 6th European Conference on Colour in Graphics, Imaging, and Vision* (pp. 321–324). Red Hook, NY: Curran Associates, Inc.
- Buchsbaum, G. A. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310, 1–26.
- Cochran, W. W., Mouritsen, H., & Wikelski, M. (2004). Migrating songbirds recalibrate their magnetic compass daily from twilight cues. *Science*, 304(5669), 405–408.
- Cornelissen, F. W., & Brenner, E. (1995). Simultaneous colour constancy revisited: An analysis of viewing strategies. *Vision Research*, 35, 2431–2448.
- Craven, B. J., & Foster, D. H. (1992). An operational approach to colour constancy. *Vision Research*, 32(7), 1359–1366.
- Crichton, S., Pearce, B., Mackiewicz, M., Finlayson, G., & Hurlbert, A. (2012). The illumination correction bias of the human visual system. *Journal of Vision*, 12(9):64, <http://www.journalofvision.org/content/12/9/64>, doi:10.1167/12.9.64. [Abstract]
- Daugirdiene, A., Murray, I. J., Vaitkevicius, H., & Kulikowski, J. (2006). Cone contrast computations: Physical versus perceived background and colour constancy. *Spatial Vision*, 19, 173–192.
- Delahunt, P. B., & Brainard, D. H. (2004b). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision*, 4(2):1, 57–81, <http://www.journalofvision.org/content/4/2/1>, doi:10.1167/4.2.1. [PubMed] [Article]
- D’Zmura, M., & Lennie, P. (1986). Mechanisms of colour constancy. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 3(10), 1662–1672.
- Epstein, W. (1973). The process of “taking-into-account” in visual perception. *Perception*, 2(3), 267–285.
- Endler, J. A. (1993). Some general comments on the evolution and design of animal communication systems. *Philosophical Transactions of the Royal Society Biological Sciences*, 340, 215–225.
- Foster, D. H. (2011). Color constancy. *Vision Research*, 51, 674–700.
- Foster, D. H., & Nascimento, S. M. (1994). Relational colour constancy from invariant cone-excitation ratios. *Proceedings of the Royal Society Biological Sciences*, 22, 115–121.
- Granzier, J., & Gegenfurtner, K. R. (2012). Effects of memory colour on colour constancy for unknown coloured objects. *iPerception*, 3(3), 190–215.
- Granzier, J., Toscani, M., & Gegenfurtner, K. R. (2012). The role of eye movements in chromatic induction. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 29, A356–A368.
- Granzier, J. J., Smeets, J. B. J., and Brenner, E. (2006). A direct test of the “grey world hypothesis”: A comparison of different matching methods. In *Third European Conference on Color in Graphics, Imaging, and Vision*, 131–136.
- Granzier, J. J. M., Brenner, E., & Smeets, J. B. J. (2009a). Can illumination estimates provide the basis for color constancy? *Journal of Vision*, 9(3):18, 1–11, <http://www.journalofvision.org/content/9/3/18>, doi:10.1167/9.3.18. [PubMed] [Article]
- Granzier, J. J. M., Brenner, E., & Smeets, J. B. J. (2009b). Reliable identification by color under natural conditions. *Journal of Vision*, 9(1):39, 1–8, <http://www.journalofvision.org/content/9/1/39>, doi:10.1167/9.1.39. [PubMed] [Article]
- Granzier, J. J. M., Nijboer, T. C. W., Smeets, J. B. J., & Brenner, E. (2005). Does realistic rendering of a gradient in illumination increase chromatic induction? *AIC Colour '05—10th Congress of the International Colour Association*, 227–230.
- Granzier, J. J. M., Smeets, J. B. J., & Brenner, E. (2012). Suggesting that the illumination differs between two scenes does not enhance color constancy. In *Proceedings of the 6th European Conference on Colour in Graphics, Imaging and Vision*, 325–330.
- Granzier, J. J. M., Vergne, R., & Gegenfurtner, K. R. (2013). The effects of surface gloss and roughness on color constancy for real 3D objects. In press.
- Hedrich, M., Bloj, M., & Ruppertsberg, A. I. (2009). Color constancy improves for real 3D objects. *Journal of Vision*, 9(4):16, 1–16, <http://www.journalofvision.org/content/9/4/16>, doi:10.1167/9.4.16. [PubMed] [Article]
- Henderson, S. T. (1977). *Daylight and its spectrum* (2nd ed). Bristol, UK: Adam Hilger Ltd.
- Hernandez-Andres, J., Romero, J., Nieves, J. L., & Lee, R. L. (2001). Color and spectral analysis of

- daylight in southern Europe. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 18, 1325–1335.
- Hering, B. (2007). Seasonal changes in dominating hues in landscapes close to nature. *Proceedings of the International Conference on Color Harmony*. Budapest, Hungary: Hungarian Academy of Sciences.
- Ishihara, S. (1969). *Tests for colour blindness: 38 plate edition*. Tokyo, Japan: Kanehara Shuppan Co. Ltd.
- Jameson, D., & Hurvich, L. M. (1989). Essay concerning color constancy. *Annual Review of Psychology*, 40, 1–22.
- Judd, D. B., MacAdam, D. L., & Wyszecki, G. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America*, 54, 1031–1040.
- Judd, D. B. (1951). Report on U.S. Secretariat Committee on colorimetry and artificial daylight. In *Proceedings of the Twelfth Session of the CIE, Stockholm, Sweden* (p. 11). Paris: Bureau Central de al CIE.
- Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt, Brace.
- Koenderink, J. J., Pont, S. C., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2007). The visual light field. *Perception*, 36, 1595–1610.
- Land, E., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, 61, 1–11.
- Lee, H. C. (1990). A computational model for opponent color encoding. *Advances Printing of conference summaries, SPSE's 43rd Annual Conference, Rochester, New York, 1990*, 178–181.
- Lee, R. L. (1994). Twilight and daytime colors of the clear sky. *Applied Optics*, 33, 4629–4638.
- Lee, R. L. Jr., & Hernández-Andrés, J. (2005a). Short-term variability of overcast brightness. *Applied Optics*, 44, 5704–5711.
- Lee, R. L. Jr., & Hernández-Andrés, J. (2005b). Colors of the daytime overcast sky. *Applied optics*, 44, 5712–5722.
- MacLeod, D. I. (2003). New trends in color perception. *Trends in Cognitive Science*, 7(3), 97–99.
- Mausfeld, R. (1998). Colour perception: From Grassman codes to a dual code for object and illuminant colours. In W. Backhaus, R. Kliegl, & J. S. Werner (Eds.), *Color vision: Perspective from different disciplines* (pp. 219–250). Berlin, Germany: Walter De Gruyter.
- McCamy, C. S., Marcus, H., & Davidson, J. G. (1976). A color rendition chart. *Journal of Applied Photographic Engineering*, 2, 95–99.
- Mollon, J. D. (2006). Monge: The Verriest lecture, Lyon, July 2005. *Visual Neuroscience*, 23, 297–309.
- Narasimhan, S. G., Wang, C., & Nayar, S. K. (2002). All the images of an outdoor scene. *European Conference on Computer Vision (ECCV)*, Vol. III (pp. 148–162), May 2002.
- Ruttiger, L., Mayser, H., Serey, L., & Sharpe, L. T. (2001). The color constancy of the red–green color blind. *Color Research and Application*, 26, 209–213.
- Shepard, R. N. (1992). The perceptual organization of colors: An adaptation to regularities of the terrestrial world? In J. Barkow, L. Cosmides, & J. Tooby, (Eds.), *The adapted mind* (pp. 495–532). Oxford, UK: Oxford University Press.
- Schirillo, J. A. (2013). We infer light in space. *Psychonomical Bulletin Review*, 20(5), 905–915.
- Smithson, H. E. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society Biological Sciences*, 360, 1329–1346.
- Teller, S., Antone, M., Bodnar, Z., Bosse, M., Coorg, S., Jethwa, M., & Master, N. (2001). Calibrated registered images of an extended urban area. *International Journal of Computer Vision*, 53(1), 93–107.
- Van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. In *Proceedings of the Royal Society of London B. Biological Sciences*, 265, 359–366.
- Von Kries, J. (1905). Die Gesichts-empfindungen. In W. Nagel, (Ed.), *Handbuch der Physiologie des Menschen 3* (pp. 109–281). Braunschweig, Germany: Vieweg und Sohn: Physiologie der Sinne.
- Walsh, J. W. T. (1961). *The science of daylight*. London: MacDonald and Company.
- Webster, M. A., & Mollon, J. D. (1997). Adaptation and the color statistics of natural images. *Vision Research*, 37, 3283–3298.
- Webster, M., Mizokami, Y., & Webster, S. (2007). Monthal variations in the color statistics of natural images. *Network: Computation in Neural Systems*, 18, 213–233.
- World Medical Association. 2004. *WMA Declaration of Helsinki—Ethical principles for medical research involving human subjects*. Retrieved from <http://www.wma.net/en/30publications/10policies/b3/>
- Wyszecki, G., & Stiles, W.S. (1982). *Color science*

concepts and methods, quantitative data and formulae. Hoboken, NJ: Wiley.

Yendrikhovskij, S. N. (2001). Computing color categories from statistics of natural images. *Journal of Imaging Science Technology*, 45, 409–417.

Zaidi, Q. (1998). Identification of illuminant and object

colors: Heuristic-based algorithms. *Journal of the Optical Society of America, A. Optics, Image Science, and Vision*, 157, 1767–1776.

Zaidi, Q. (2001). Color constancy in a rough world. *Color Research & Application*, 26, 192–200.