Illumination assumptions account for individual differences in the perceptual interpretation of a profoundly ambiguous stimulus in the color domain: "The dress"

Pascal Wallisch

Department of Psychology, New York University, New York, NY, USA

There has been considerable interest in a stimulus ("the dress") that yields starkly divergent subjective color percepts between observers. It has been proposed that individual differences in the subjective interpretation of this stimulus are due to the different assumptions that individuals make about how the dress was illuminated. In this study, we address this possible explanation empirically by reporting on data from ∼13,000 observers who were surveyed online. We show that assumptions about the illumination of the dress—i.e., whether the stimulus was illuminated by natural or artificial light or whether it was in a shadow—strongly affects the subjective interpretation of observers, compared to demographic factors, such as age or gender, which have a relatively smaller influence. We interpret these findings in a Bayesian framework by also showing that prior exposure to long- or short-wavelength lights due to circadian type shapes the subjective experience of the dress stimulus in theoretically expected ways.

Introduction

Perception is inherently idiosyncratic. Encased in a bony shell that is completely opaque, the brain receives information about the outside world through sensory organs—in the case of visual perception, the eyes. In primates, the information coming from the retinae is interpreted by a large network of highly interconnected visual areas (Felleman & Van Essen, 1991; Wallisch & Movshon, 2008). Thus, it should not come as a surprise that the interpretation of this information by the visual system of any given individual—infused with plenty of extraretinal information (Yeshurun & Carrasco, 1998; Changizi & Hall, 2001; Bar et al., 2006; Anderson, Siegel, Bliss-Moreau, & Barrett, 2011)—does not necessarily correspond veridically to conditions in the external world or the interpretation of other observers.

In terms of behavioral effects, this has been known for a long time in the domains of object perception, such as with ambiguous figures, like the Necker cube (Necker, 1832) or the duck/rabbit illusion (Brugger, 1999); the perception of luminance as with the Hermann grid and variations thereof (Hermann, 1870; Ninio & Stevens, 2000); and motion as in the case of the reconstruction of the three-dimensional structure of objects from motion (Wallach & O'Connell, 1953).

In the domain of color vision, we know of effects that can result in large discrepancies between the subjective percept and what one would expect from the wavelengths present in the stimulus. For instance, the local—patterned—background of a colored line can strongly shift the color appearance of that line (Monnier & Shevell, 2003). More fundamentally, the brain is always solving an inverse problem when deducing the color of an object from the spectral information given by the eyes. Specifically, it has to discount the wavelengths of the illuminant on an ongoing basis—a phenomenon known as "color constancy"—which can lead to dramatic shifts in the perceived color of suitably designed stimuli (Lotto & Purves, 1999). However, these shifts away from veridical—dramatic as they might be—seem to happen in the same direction for all observers without color vision deficiencies. In other words, visual illusions with starkly divergent interpretations between observers were known in the domain of object, motion, or luminance perception but not color perception. It can be argued that such stimuli were known in the color literature, e.g., Bloj, Kersten, and Hurlbert (1999), but it can be contended that this does not constitute a genuine bistable stimulus as one typically requires the retinal image not to change for those, which is not the case for Bloj et al.

Regardless, this situation changed in February 2015, when an image of "the dress" was posted online, where it was met with immediate and considerable interest,

Citation: Wallisch, P. (2017). Illumination assumptions account for individual differences in the perceptual interpretation of a profoundly ambiguous stimulus in the color domain: "The dress". *Journal of Vision*, *17*(4):5, 1–14, doi:10.1167/17.4.5.

doi: 10.1167/17.4.5 Received September 30, 2016; published April 7, 2017 ISSN 1534-7362

both within the vision science community and outside of it.

The image itself was taken by Cecilia Bleasdale on February 7 at 3:30 in the afternoon with a Galaxy S3 mini cellphone in a shop in England on a cold winter day with dusk falling (Cecilia Bleasdale, personal communication, May 15, 2015). It depicts a black-and-blue dress that she intended to wear to her daughter's wedding later in the month.

The reason for the global interest in this image was that different observers spontaneously came to radically divergent interpretations regarding the color of this image even when viewing it on the same screen.

Vision scientists were quick to investigate this phenomenon and confirmed that the divergence effect is real: People really do settle into stable but drastically different perceptual clusters, and the reported color percepts do not seem to be an artifact of categorical color labels (Lafer-Sousa, Hermann, & Conway, 2015). Moreover, it could be shown that there is a special potential of confusion between these colors; i.e., inverted colors do not produce ambiguous percepts (Winkler, Spillmann, Werner, & Webster, 2015). The effect seems to be most prominent for colors that follow the spectral distribution of natural daylight from blue to yellow, i.e., the "daylight axis" (Gegenfurtner, Bloj, & Toscani, 2015).

Of course, after establishing that there is something genuinely confusing about the visual interpretation of the color of this specific image, the critical question immediately arises whether we can predict what a given individual observer will see and, if we can do so, what this tells us about the determinants of subjective color perception. Early reports suggested a modest effect of gender with women being more likely to see white and gold (Lafer-Sousa et al., 2015) but for unclear reasons. Efforts to predict the color percept from genetic markers were unsuccessful even in very large samples (Sathirapongsasuti, 2015), and even monozygotic twins seem to exhibit a surprisingly high degree of perceptual variability (Mount, 2015), making a purely genetic basis of this phenomenon unlikely. Recent research suggests that the proportion of variance accounted for by genetics is on the order of 33% with the remainder due to environmental factors (Mahroo et al., 2017).

In addition, there have been attempts to explain individual variation in the perceptual experience of the dress in terms of neuroscience and psychology. In principle, this variation could be due to low-level (individual differences in the sensory apparatus itself) or high-level (individual differences in feedback to the visual system) factors or both. One attempt to explain individual differences in terms of low-level effects was provided by Rabin, Houser, Talbert, and Patel (2016). They measured the macular pigment optical density (MPOD) and related it to the subjective experience of the color of the dress stimulus. They found that participants with denser MPOD are more likely to perceive the dress as white and gold. This is an interesting observation, but the reported effect size in this study is modest, so whereas this effect might be contributing to intersubjective variability, it would be surprising if it could account for the full range of diverse experiences reported by observers when viewing the dress.

An attempt at a high-level explanation of the intersubjective variability of color percepts in response to the dress was put forward by Schlaffke et al. (2015). They recorded brain activity in response to the dress stimulus with fMRI and showed that participants who experienced the dress as white and gold were more likely to exhibit activity in higher order cortical areas, suggesting that feedback from these areas—instantiated by additional neural activity—overrides the veridical perception of the dress as black and blue. Whereas this explanation is intriguing, we are concerned about a category error in the interpretation of these findings. The original dress is in fact black and blue, but this information was not available to the observers in the scanner. The image of the dress that these observers were asked to judge is much more ambiguous, and it is unclear what a veridical percept of the—overexposed—image would be.

Thus, both of these attempts at explaining the subjective variability of the dress by high- and low-level mechanisms are certainly thought provoking but ultimately not fully convincing. One concern that is common to both studies is that the actual mechanism that would compel one kind of participant to see the dress one way but not another way remains vague. A second concern is that both studies are relatively low powered.

Thus, we want to consider an alternative explanation here: one that is perhaps more compelling as it is backed by a large degree of literature on how we think color vision works normally as well as one that is specific enough to make clear predictions. This explanation was originally put forward by Bevil Conway in a *Wired* article (Rogers, 2015) and then again in Lafer-Sousa et al. (2015). It posits that the visual system is constantly engaged in color correction as it has to recalibrate itself on an ongoing basis in order to discount the illuminant and achieve color constancy to preserve object identity in the face of changing illuminations (D'Zmura & Lennie, 1986). Note that this is something the visual system would have to do even before the advent of modern illumination sources; the spectral composition of daylight varies throughout the day, sometimes objects are in the shade, and so on. Therefore, we know that the visual system is trying to achieve color constancy and has a mechanism in place to bring it about. But what would such a system do if it encountered a stimulus for which the illumination is ill defined or ambiguous? We believe that the image of the dress does qualify as the upper part of the image implies relatively short-wavelength illumination by daylight and relatively long-wavelength artificial illumination in the lower part of the image. In other words, observers can be expected to rely on their assumptions about illumination to guide which illuminant color constancy mechanisms will try to discount. If that were the case, we would expect there to be diversity in the subjective experience of such a stimulus as we cannot expect everyone to have the same illumination assumptions. Of course, we see this kind of diverse interpretation in the case of the dress stimulus. Moreover, we can make much more specific predictions. Those who assume a relatively long-wavelength illuminant—such as from incandescent light—could be expected to see the dress as black and blue as they would subtract long wavelengths from the image in order to arrive at a relatively shorter wavelength percept of the object, and consequently, the dress stimulus would appear yellowish. Conversely, those who assume a relatively short-wavelength illuminant, such as daylight, could be expected to see the dress as white and gold as they would subtract short wavelengths from the image in order to arrive at a relatively longer wavelength percept of the object; consequently, the dress stimulus would appear bluish. The purpose of the current study is to investigate this possible explanation directly and empirically.

Method

To answer this question—whether the subjective interpretation of the dress stimulus can be accounted for by individual assumptions about illumination—we employed the following procedures.

Participants

We logged data from a total of 13,417 participants from around the world who responded to online surveys. These data were logged in two runs, one starting in March 2015 and one in March 2016, yielding data from 8,084 (Run 1) and 5,333 participants (Run 2), respectively. The bulk of the data from each run (>97%) was logged within the first month of publication of the recruitment links. Given that the data stems from an online sample that can be expected to be noisy, it is important that it is sufficiently large to be adequately powered (Wallisch, 2015a). Participants were recruited by social media (tweets with the survey link) and media; the link to the survey was posted at the end of popular science pieces that discussed the reasons the dress phenomenon is interesting to vision science as well as the societal relevance of vision research on Slate.com and Elsevier Connect.

Survey

All participants were asked to answer a number of question about their subjective experience of the dress stimulus, such as what colors they saw when they first saw the dress stimulus and whether their percept switched since then. Response options by participants were categorical in terms of the most commonly reported percepts (white/gold, black/blue, white/blue, white/black, rapidly switching, and "other"). In the survey for Run 1, we also asked whether they thought the dress was in a shadow, whether they knew the true color of the dress when they first saw it, and the time and day when they first saw the dress as well as the device they saw it on and whether it was inside or outside. In addition, we asked a number of questions about the participants themselves: whether they grew up in the city or the country or whether they currently live in city or country, their self-identified circadian type, and whether they spend most of their time inside or outside. In the survey for Run 2, we again asked about the percept of the participants when they first saw the dress as well as the factors that were deemed significant in the analysis of the data from Run 1, namely whether they believed the dress to be in a shadow, whether they knew its true color, and their circadian type. We also asked a question that was not asked in Run 1, namely whether they believed the dress was illuminated by artificial or natural lighting. Finally, we asked demographic questions such as age and gender in both runs. These surveys were deployed online on Surveymonkey.com. All procedures were approved by the New York University Institutional Review Board (UCAIHS).

Data analysis

All data were analyzed with MATLAB (Mathworks, Inc., Sherborn, MA). Unless noted otherwise, we first performed a chi-squared test—comparing expected and observed counts of the percept—in response to any given question to see if there were statistically significant differences in the propensity to see the dress in a certain way. Then, if this test yielded statistical significance, we prepared a corresponding figure. The error bars in these figures represent the standard error of the mean, which was calculated by resampling methods (Efron & Tibshirani, 1994). Briefly, we sampled—with replacement—100,000 times from the

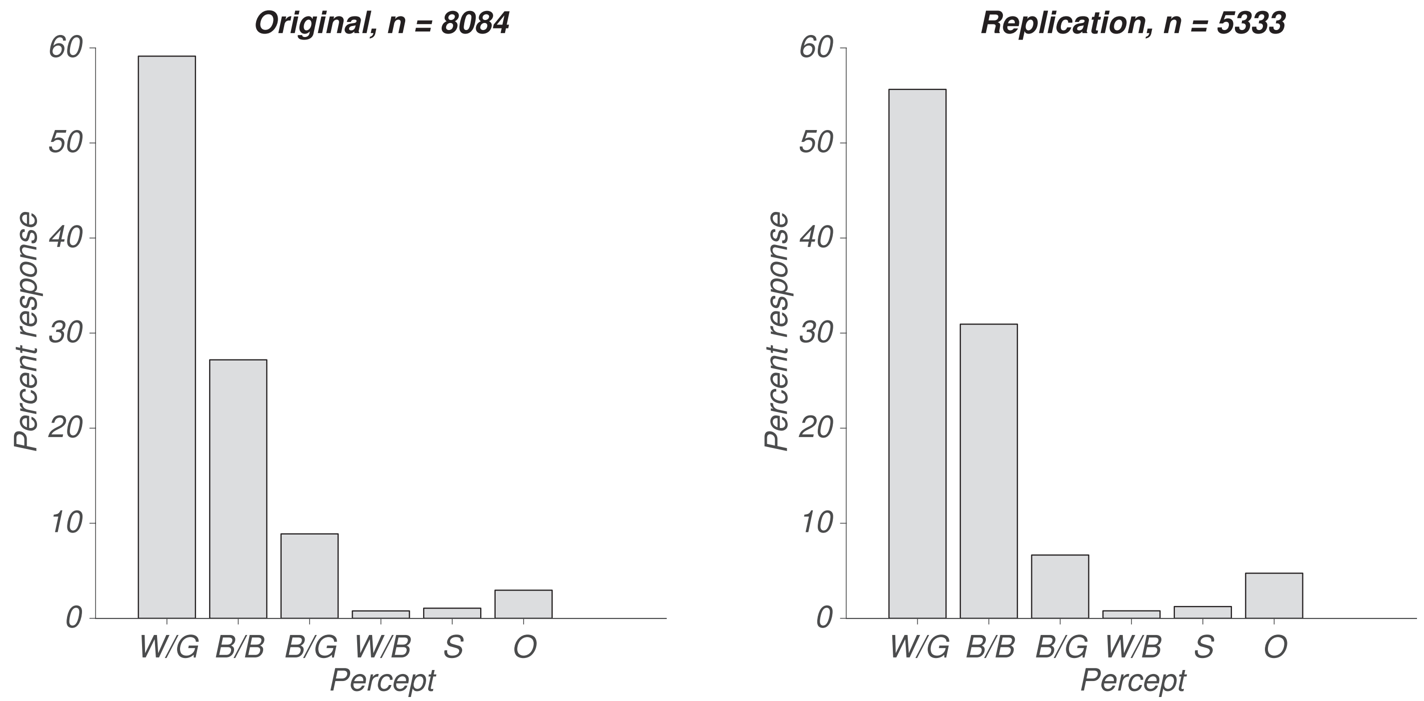

Figure 1. Color percept of participants upon first seeing the dress stimulus. The *x*-axis represents the categorical response. W/G = white and gold, B/B = blue and black, B/G = blue and gold, B/W = black and white, S = switching or bistable, O = other. The *y*-axis represents the proportion of participants that report seeing the dress stimulus in this way, in percentages. Left panel: data from Run 1; right panel: data from Run 2.

responses of any given group of participants (e.g., women), then calculated the standard deviation of these 100,000 sample means. This standard deviation corresponds to the estimated standard error of the mean. To guard against issues of multiple comparison (we performed 17 chi-squared tests in this manuscript), we dropped the significance level to 0.01.

Results

Proportions of participants who see the dress stimulus in a certain way

We asked participants to report their color percept upon first seeing the dress image. In our sample, for Run 1, 59.13% ($n = 4{,}780$) of participants reported seeing the dress as white and gold, 27.19% ($n = 2{,}198$) of participants reported seeing the dress as black and blue, 8.88% ($n = 718$) of participants reported seeing the dress as blue and gold, and 0.78% ($n = 63$) of participants reported seeing the dress as white and black whereas 1.06% ($n = 68$) of participants reported seeing the dress as switching or ambiguous, and 2.96% ($n = 239$) reported an "other" or idiosyncratic percept. The corresponding numbers from Run 2 are 55.63% ($n = 2{,}967$), 30.94% ($n = 1{,}650$), 6.66% ($n = 355$), 0.79% ($n = 42$), 1.23% ($n = 66$), and 4.74% ($n = 243$), respectively. As you can see, there was a slight decrease in reporting the percept as white and gold and blue and gold and a slight increase in reporting the percept as black and blue and other from Run 1 to Run 2, but overall, the proportions are rather stable—as a matter of fact, none of the relative ranks has changed between runs (see Figure 1).

We consider the observation that there is dramatic diversity in the color experiences of the dress stimulus by observers to be the "dress effect." It is what we will attempt to explain by linking it to illumination assumptions in the rest of the manuscript. Reassuringly, these proportions are in line with reports from other large-scale, online efforts to poll the color responses immediately after the dress phenomenon broke (Holderness, 2015; Sathirapongsasuti, 2015) with a majority of participants reporting seeing the dress as white and gold. Thus, we will from now on consider "white and gold" the canonical percept and report the proportion of participants who report seeing the dress as white and gold versus all other possibilities.

Summary of effects

Table 1 summarizes the effects we found in both runs. We will discuss them in detail below.

	Run 1			Run 2		
Effect	χ^2	*df*	*p*	χ^2	*df*	*p*
Shadow	**326.90**	**2**	**<0.01**	**511.69**	**2**	**<0.01**
Lighting	**-**	**-**	**-**	**147.56**	**2**	**<0.01**
Switching	**99.63**	**2**	**<0.01**	**-**	**-**	**-**
True color	**27.02**	**2**	**<0.01**	**13.75**	**2**	**<0.01**
Circadian	**11.39**	**3**	**<0.01**	**90.56**	**3**	**<0.01**
Gender	4.57	1	<0.05	2.48	1	>0.05
Time seen	4.53	5	>0.05	-	-	-
Day seen	4.43	5	>0.05	-	-	-
Seen inside	4.03	2	>0.05	-	-	-
Screen type	3.61	3	>0.05	-	-	-
Time outside	2.88	2	>0.05	-	-	-
Growing up	2.51	2	>0.05	-	-	-
Living now	1.92	2	>0.05	-	-	-

Table 1. Effects and associated statistical parameters for data from Runs 1 (original) and 2 (replication). *Notes*: Effects are ordered by absolute chi-squared value. Effects that are significant at $p < 0.01$ are bolded.

Assumptions about light that influence the perception of the dress stimulus

We asked participants to report whether they believed that the dress was in a shadow or not. In our sample, 2,151 participants did believe this to be the case, 3,222 participants did not believe that the dress was in a shadow, and 2,711 participants from Run 1 indicated that they were not sure. The corresponding numbers from Run 2 are 1,488, 1,984, and 1,861, respectively. We related these beliefs to the color experiences that participants reported (see Figure 2).

As you can see, beliefs about whether the dress was in a shadow or not did strongly modulate the perceptual experience of the dress with a modulation range of between 20% (as suggested by Run 1) and 40% (as suggested by Run 2). This result constitutes a conceptual replication of Chetverikov and Ivanchei (2016), who showed that the perceived direction of illumination of the dress is related to percept. This result is also in line with theoretical predictions from the general hypothesis that illumination assumptions drive the dress effect: Shadows overrepresent short wavelengths. In other words, shadows appear bluish. If someone assumes that the dress was in a shadow, color constancy mechanisms could be expected to discount the effect of the shadow, rendering the conscious appearance of the dress more yellowish, relative to someone who does not believe this to be the case.

Where might such beliefs come from? It would be hard to ascertain why someone assumes that something is in a shadow or not. In general, one can expect that people assume things to be more likely that they have encountered more often in their past in a Bayesian

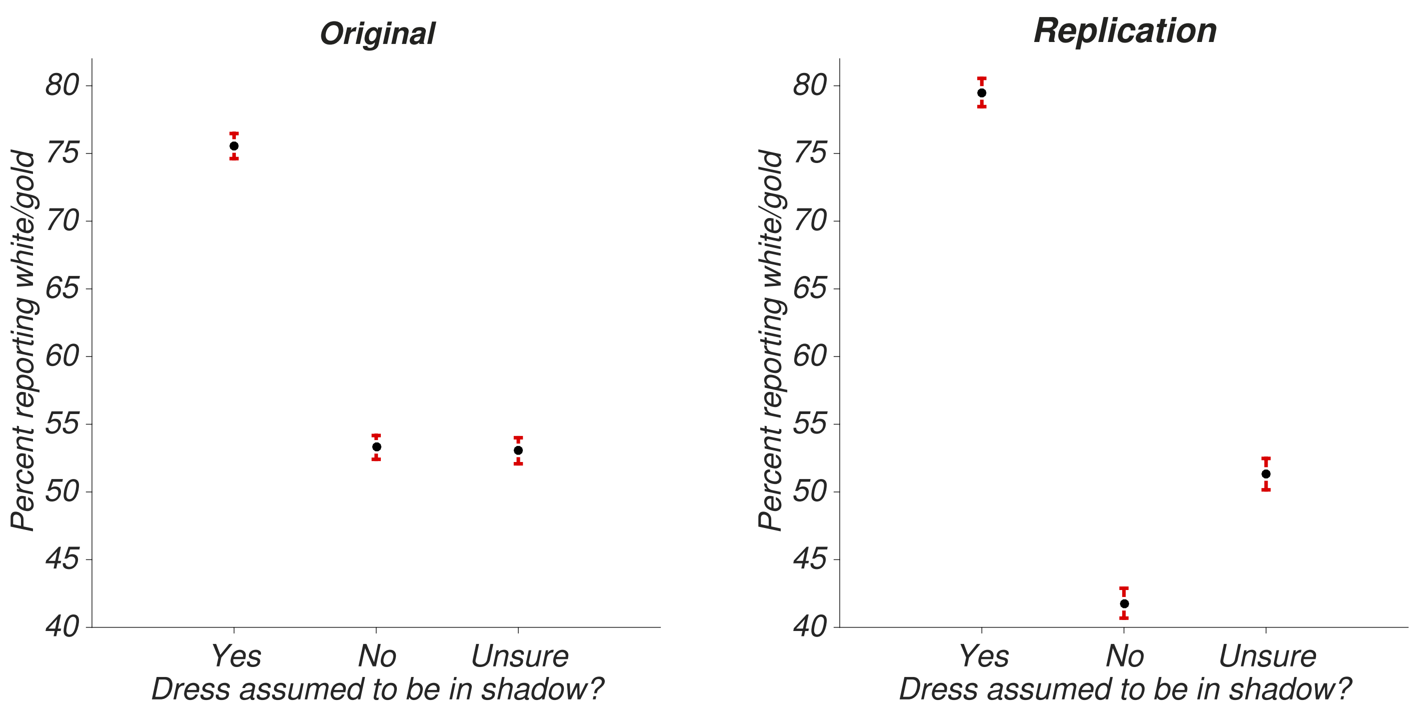

Figure 2. Color percept of the dress stimulus as a function of the belief as to whether the dress was in a shadow. The *x*-axis represents different groups of participants: those who believed the dress to be in a shadow, those who believed the dress was not in a shadow, and those who were unsure. The *y*-axis represents the proportion of participants reporting seeing the dress stimulus as white and gold. Black dots represent the mean response of a given group of participants; error bars represent the *SEM*. Left panel: data from Run 1; right panel: data from Run 2.

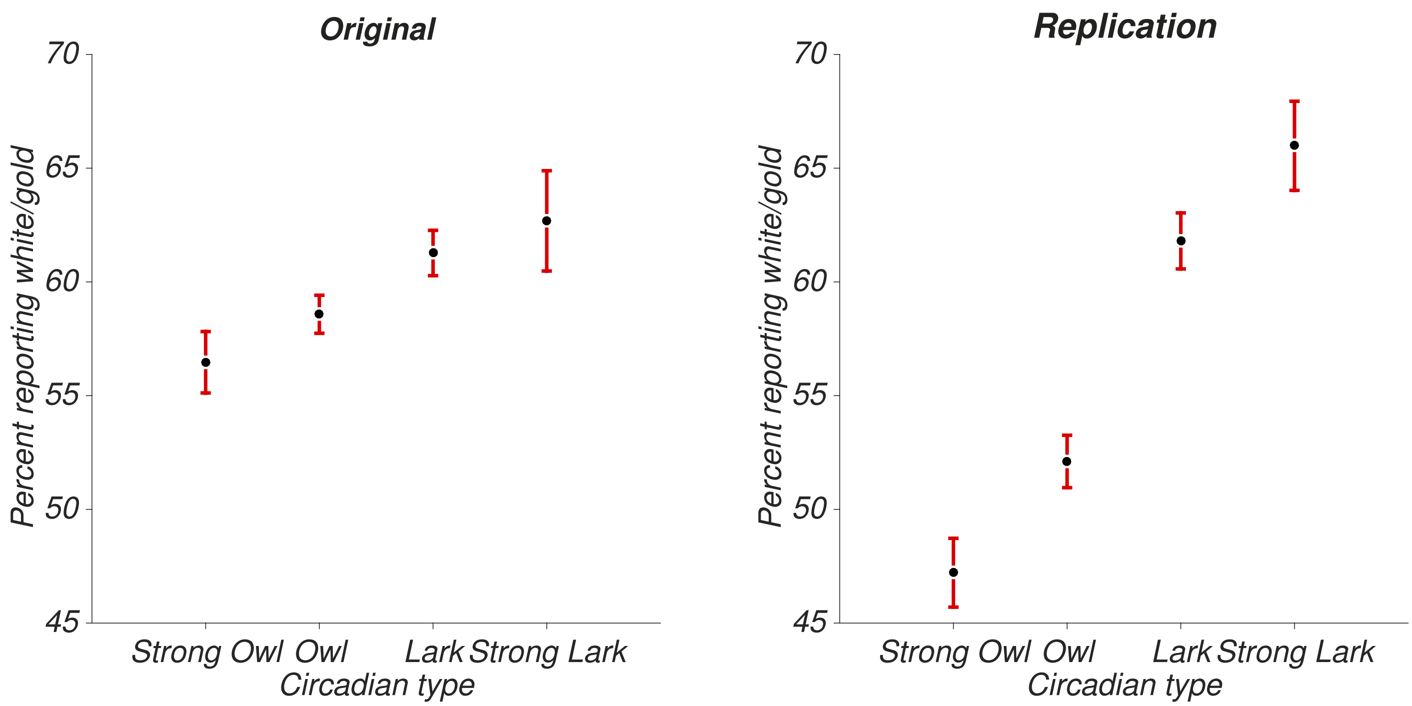

Figure 3. Color percept of the dress stimulus as a function of circadian type. The *x*-axis represents self-identified circadian type. The *y*-axis represents the proportion of participants reporting seeing the dress stimulus as white and gold. Black dots represent the mean response of a given group of participants; error bars represent the *SEM*. Left panel: data from Run 1; right panel: data from Run 2.

fashion. For instance, slower speeds are much more commonly encountered than faster speeds, so if someone is exposed to moving stimuli under observation conditions with some uncertainty, they will take into account that slower speeds are more likely than faster ones and downgrade their speed estimates accordingly (Stocker & Simoncelli, 2006). In the color domain, one would expect that someone who has encountered more short-wavelength light to assume a short-wavelength (or bluish) illumination if there is uncertainty about the nature of the illuminant. In contrast, someone who has encountered more long-wavelength lights in their past would be more likely to assume a long-wavelength (or yellowish) illuminant in the same situation. Those are specific predictions given access to someone's illumination prior, but how would one go about estimating someone's lifetime exposure to lights of different spectral content? One way to go about this is to assume that people who spend more time in daylight will spend more time being exposed to natural light whereas those who spend more waking time at night will be more exposed to artificial—until recently, usually incandescent—light. Fortuitously for us, there is quite some diversity in terms of the human propensity to get up at sunrise and go to bed at sunset ("larks") versus those who are phase-shifted, getting up late and staying up later ("owls"). This circadian type is largely stable throughout the lifetime after around age 25 and seems to be strongly determined genetically (Gibertini, Graham, & Cook, 1999; Adan et al., 2012; Pegoraro et al., 2015). In addition, natural sunlight has different spectral properties than incandescent artificial light: The spectral power of incandescent light is shifted toward longer wavelengths (Smith, 2016). In other words, it seems reasonable to assume that larks have been exposed to more bluish sunlight than owls, who are exposed to more yellowish incandescent light. If so, we could use circadian type as a proxy for illumination priors and would predict that larks tend to see the dress stimulus as white and gold as they discount for the assumed bluish illumination whereas owls can be expected to see the dress stimulus as blue and black as they discount an assumed yellowish illuminant. With this rationale and these predictions in place, we asked the participants about their self-identified circadian type (lark = gets up early, goes to bed early, and feels best in the morning vs. owl = likes to sleep in and stay up and feels best at night).

For Run 1, 1,353 participants identified as "strong owls," 3,486 participants as "owls," 2,409 participants as "larks," and 477 participants as "strong larks." The corresponding numbers from Run 2 are 1,095, 1,875, 1,576, and 585, respectively. We related these beliefs to the color experiences that participants reported (see Figure 3).

As you can see, self-identified circadian type is associated with how the dress is perceived in both runs in a direction consistent with the prediction from the

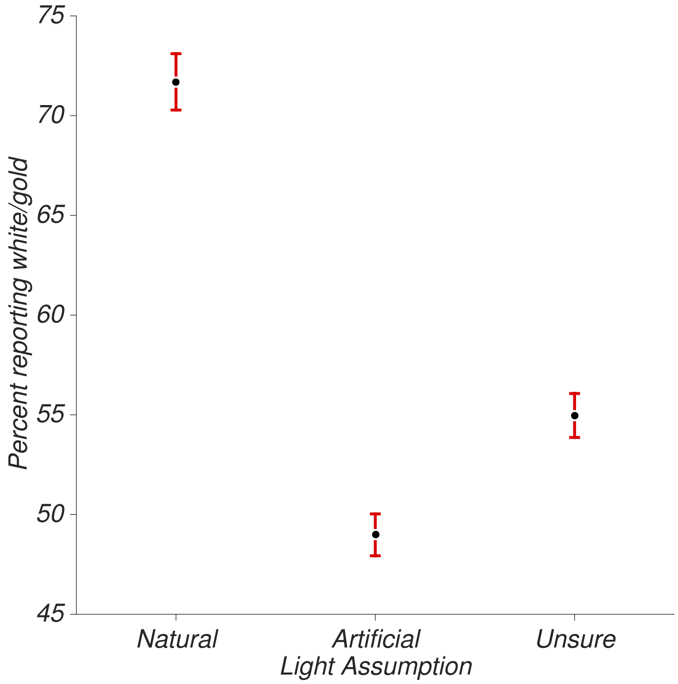

Figure 4. Color percept of the dress stimulus as a function of beliefs about illumination type from Run 2. The *x*-axis represents different groups of participants: those who believed the dress to be illuminated by natural light, those who believed the dress to be illuminated by artificial light, and those who were unsure. The *y*-axis represents the proportion of participants reporting seeing the dress stimulus as white and gold. Black dots represent the mean response of a given group of participants; error bars represent the *SEM*.

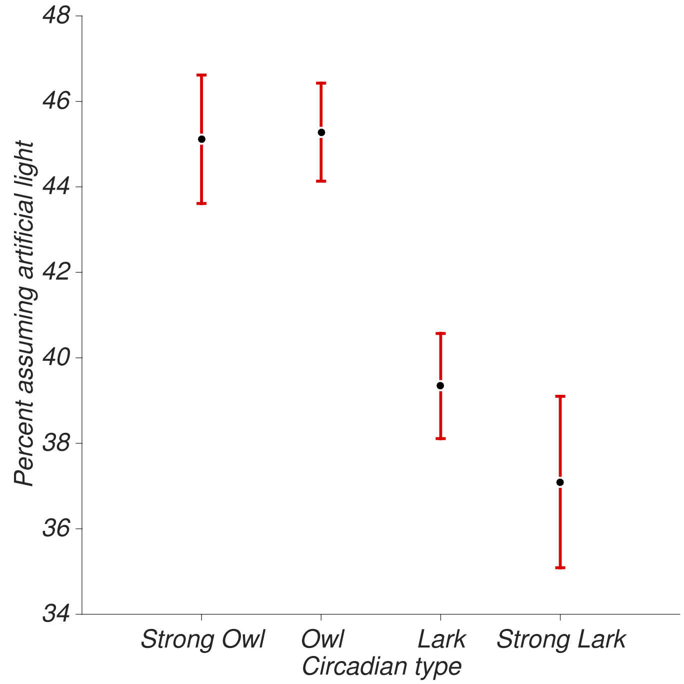

Figure 5. Proportion of participants from Run 2 assuming artificial illumination of the dress as a function of circadian type. The *x*-axis represents self-identified circadian type. The *y*-axis represents the proportion of participants reporting to have assumed an artificial light source. Black dots represent the mean response of a given group of participants; error bars represent the *SEM*.

hypothesis that illumination assumptions explain the dress effect and in a dose-dependent fashion.

We asked the study participants from Run 2 to directly report their assumptions about the nature of the illumination. The image itself is ambiguous as to whether the dress was photographed outside and being illuminated by natural light or inside and illuminated by artificial light. In our sample, 2,285 participants believed the dress to be illuminated by artificial light, 1,021 participants by natural light, and 2,054 participants were not sure. We related these beliefs to the color experiences that participants reported (see Figure 4).

As you can see, assumptions about the nature of the illuminant are strongly associated with different perceptions of the dress in a direction consistent with the prediction from the hypothesis that illumination assumptions explain the dress effect.

To close the logical loop, we can now look at whether people who self-identify as owls are more likely to assume artificial light than those who self-identify as larks. If our rationale is internally consistent, they should, as they are more likely to encounter artificial illuminations at night. As you can see in Figure 5, this is in fact the case. Eliminating those who were unsure about illumination conditions, self-described owls have a higher likelihood of assuming artificial illumination in the dress stimulus ($\chi^2 = 22.43$, $df = 3$, $p < 0.01$).

Taken together, we believe that these four strong and consistent effects lend strong support to the hypothesis that assumptions about illumination underlie the dress effect.

Further evidence that color constancy mechanisms play a role in the dress effect

Color constancy effects are particularly strong when the observer knows the true color of the object (Hansen, Olkkonen, Walter, & Gegenfurtner, 2006). For instance, strawberries will continue to appear red even if illuminated by a green laser pointer because people know that strawberries are red. As the illuminant is ill defined in the picture of the dress, we would expect that it is particularly easy to discount it if people know the true color of the dress when they first see the image of the dress. This seems to be the case empirically (see Figure 6).

However, this is complicated by the fact that most participants were unaware as to the true color of the

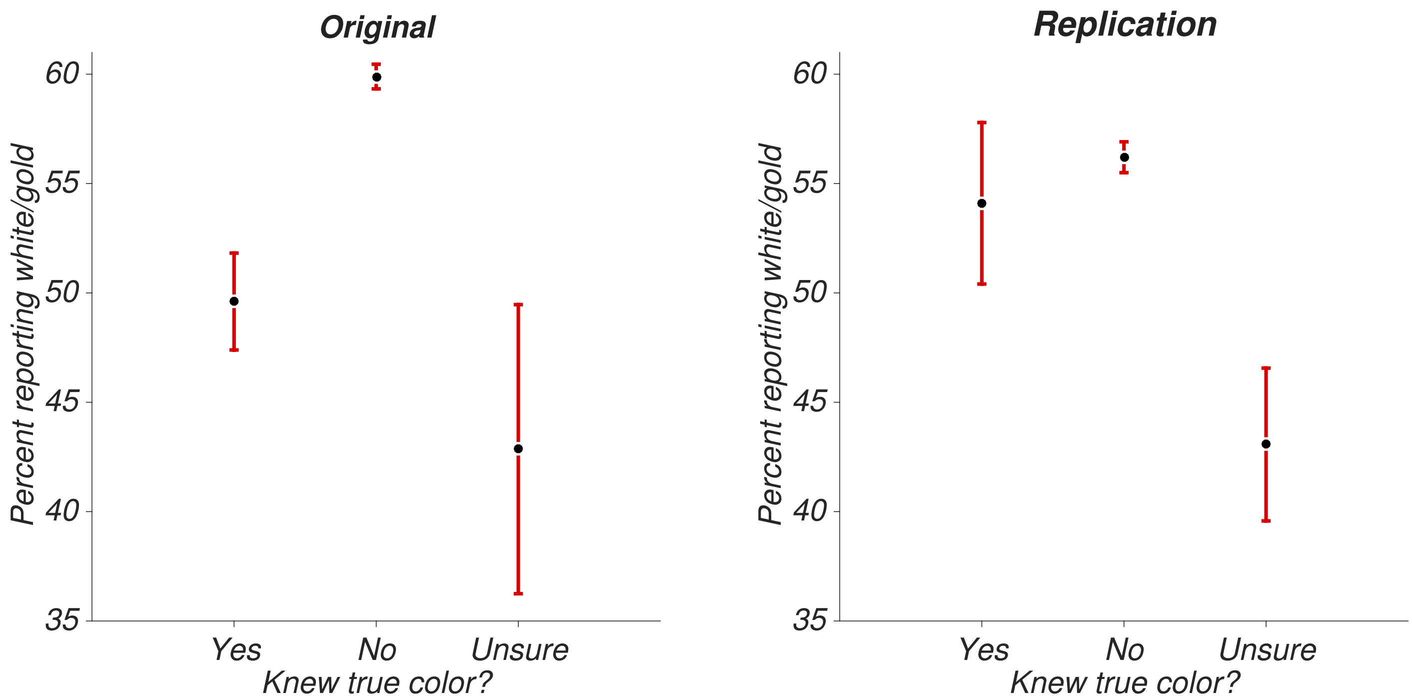

Figure 6. Color percept of the dress stimulus as a function of knowing the true color of the dress. The *x*-axis represents different groups of participants: Those who reported knowing the true color of the dress and those who did not. The *y*-axis represents the proportion of participants reporting seeing the dress stimulus as white and gold. Black dots represent the mean response of a given group of participants; error bars represent the *SEM*. Left panel: data from Run 1; right panel: data from Run 2.

dress; only a small percentage of participants reported to be aware of the true color of the dress. If they were aware of the true color, they were more likely to see it veridically as black and blue. This meshes nicely with the observation that observers were more likely to switch from a white-and-gold percept to the true color of the dress than vice versa. We asked the participants of the first run if their percept had switched since the first time they saw the dress or not. Most participants reported that the percept had not switched since first exposure whereas a minority reported that it did switch sometime between the first exposure and when taking the survey. Note that the switching probability is asymmetric; most of those who switched did so from white and gold to the true color of the dress, not vice versa (see Figure 7). This also agrees nicely with recent reports in the literature that participants who reported the dress to be white are more likely to be influenced by contextual illumination cues (Toscani, Gegenfurtner, & Doerschner, 2017).

Effects of demographic variables: Gender and age

There have been early reports that gender plays a modest but consistent role in the perception of the dress with females being slightly more likely to report seeing the dress stimulus as white and gold (Lafer-Sousa et al., 2015; Sathirapongsasuti, 2015). Thus, we asked our participants to report their gender. In Run 1, 5,069 of our participants self-identified as female, and 2,799 of our participants self-identified as male whereas 3,162 and 2,076 did so in Run 2, respectively. However, as you can see in Table 1 and Figure 8, the effect of gender on the reported percept of the dress is small and statistically unreliable.

Finally, we asked participants to report their age in both runs. The median age in our sample was 37, which is 1 year above the median age in the United States at the time the survey was taken, which makes sense as a fraction of our participants took the survey from Europe, where the median age is slightly higher. In our sample, 99.8% reported ages below 76 years of age, which is why we only graph age until 75; beyond that, sample size per age bin is too small to meaningfully represent proportions. In addition—to get stable age bins—we combine data from both runs in this analysis. The proportion of white-and-gold percepts graphed as a function of age is interesting. It starts in late adolescence around 0.5, then gradually climbs to just above 0.6, where it is stable for about 20 years. Then, the degree of white-and-gold percepts drops sharply, starting just above age 65 and dropping to about 0.3 by age 75 (see figure 10). Generally speaking, this trend also replicates the findings from Lafer-Sousa et al. (2015) and Sathirapongsasuti (2015); white-and-gold

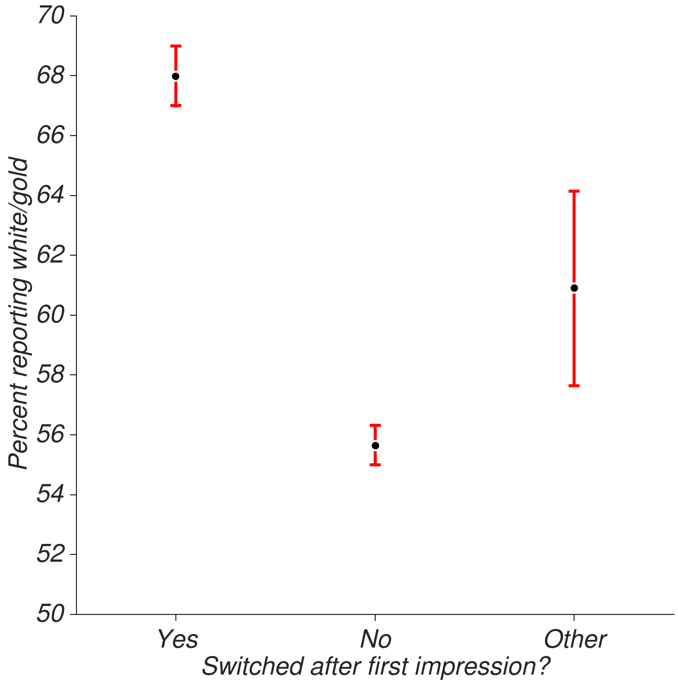

Figure 7. Color percept of the dress stimulus as a function of whether the color percept switched after first exposure from Run 1. The *x*-axis represents different groups of participants: Those who reported that their percept switched at some point after first encountering the dress stimulus, those who reported that it did not and those who provided a more complex description of their switching experience. The *y*-axis represents the proportion of participants reporting seeing the dress stimulus as white and gold. Black dots represent the mean response of a given group of participants; error bars represent the *SEM*.

percepts rise from the 20s to the 40s, then drop—and precipitously so—with advanced age.

These dynamics are interesting, but very hard to interpret. Age as a carrier variable inherently confounds many other variables that are actually causal, making it almost impossible to distinguish between possible explanations that could account for the effect of age. For instance, it is conceivable that increasing yellowing of the lens with advancing age might change how observers perceive images in general, not just the dress stimulus (Salvi, Akhtar, & Currie, 2006). Similarly, age-related changes of the optical system of the eyes might gradually dim the image with advanced age. If image brightness is associated with color percepts of the dress—brighter images being associated with white-and-gold percepts (Vemuri et al., 2016)—this could account for a reduction in black-and-blue percepts with advanced age all by itself. But there are many changes to eyes as a function of age, for instance, an increase in light scattering, which might affect the perception of the dress stimulus (Watanabe, Fujii, Nakamura, & Korenaga, 2015). However, in addition to age-related changes to the structure and function of the eye itself, one can't rule out generational effects in a cross-sectional sample like this. We really cannot dismiss the possibility that these people were exposed to different priors when they were at the same age. For instance, older people can be expected to have spent more time outdoors in their youth compared to current youth whereas incandescent lights are gradually being replaced with LED lights, which would produce differential prior exposure effects on different age cohorts. Perhaps it is as simple as the priors being weak in the young, accounting for the rise of white-and-gold percepts as participants approach middle age, and increasing insomnia in the elderly, presumably exposing them to plenty of artificial light. This last idea could be powerful. There is no telling how long it takes for priors to become established or over how much time they integrate information. It must not be a lifetime. For instance, it is conceivable that the illumination prior takes only the last year or so of light information into account. Figure 9 looks suspiciously like a graph of workforce participation rate by age (McBride, 2011). The sharp drop in white-and-gold percepts coincides roughly with the beginning of retirement age in the Western world, where most of the participants in our sample reside. Could it be that workforce participation correlates with looking at bright—and perhaps bluish—screens for a large part of the day, which might shift the spectral composition of the average light seen, and thus the light priors, toward shorter wavelengths?

However tempting it is to speculate, it is important to emphasize that we really can't distinguish between any of these potential explanations on the basis of this data alone.

Discussion

We believe we have shown that it is assumptions about illumination priors that really do matter most when determining the percept of an individual when viewing the dress stimulus for the first time.

However, there are several limitations that could potentially pose a challenge to these conclusions. First, our data were logged online, which is inherently problematic; we do not know what screens or screen settings our participants used when they viewed the dress stimulus originally. However, this lack of controlled viewing conditions might be less troubling than usual in this particular case for several reasons. First, we tried to account for the divergent phenomenological experience of our participants when first encountering the dress—whatever the viewing conditions. This divergence is the key issue at hand that made this phenomenon interesting; it neither relies on

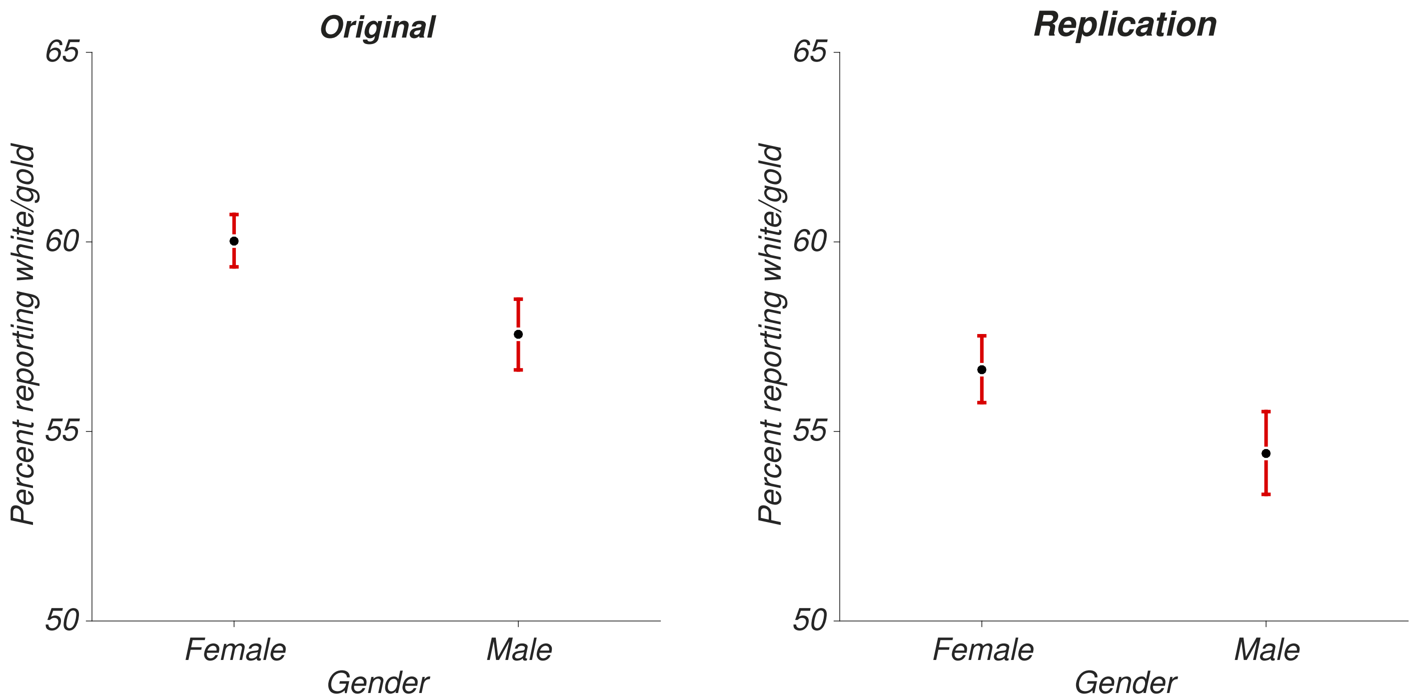

Figure 8. Color percept of the dress stimulus as a function of gender. The *x*-axis represents self-identified gender. The *y*-axis represents the proportion of participants reporting seeing the dress stimulus as white and gold. Black dots represent the mean response of a given group of participants; error bars represent the *SEM*. Left panel: data from Run 1; right panel: data from Run 2.

nor is created by particular viewing conditions. Second, the fact that it is possible to find consistent and statistically reliable responses highlights the robustness of these effects; they do not rely on carefully controlled viewing conditions unlike so many effects in vision science that do. Indeed, Chetverikov and Ivanchei (2016) show that the percept of an individual is rather stable with no statistically reliable effect of image size or device type. Third, we conceptually replicate other findings by authors who did aim to control viewing conditions, e.g., Chebichevski and Ivanchei, so we are somewhat confident that the novel findings we report are also not artifactual.

A second potential limitation is that we asked about color categories instead of doing color-matching experiments. However, previous research, e.g., Lafer-Sousa et al. (2015) or Chebichevski and Ivanchei (2016), that used color matching established that differences in the reported subjective percept of the dress cannot be attributed to differences in color labeling. Yet it would have been better to query subjective experience in a more continuous fashion. Arguably, any continuous distribution yields a bimodal one when investigated with forced choices as we did here. Then again, #thedress became a viral phenomenon not just because the subjective interpretation of this image was somewhat ambiguous, but because the reported qualia were so strikingly divergent.

Third, it is somewhat puzzling that the effects of the second run—particularly the effects of assuming the dress being in a shadow and circadian type—came out so much stronger despite being delayed by a year, relative to the first run. We can only speculate as to why that might be the case, but perhaps a parsimonious explanation could involve the fact that the dress phenomenon attracted a tremendous amount of attention in early 2015 with more than just a few people clicking on links due to curiosity and without any intention of providing usable data. In contrast, excitement about the dress had died down considerably outside of the vision science community by early 2016, which might have raised the threshold for participation. In other words, this difference might be due to a relatively higher data quality. This possibility is not implausible given the fact that we did not exclude data from any participant here; we used all of it. Finally, it is a limitation that as we asked our participants to recall their percept from memory, all of them were necessarily not naïve.

Despite these potential limitations, we are reassured by the fact that our results are consistent, high powered, and replicate all other known dress effects.

If it is the case that subjective illumination priors primarily drive the dress effect, it might have wide-ranging implications. Up until now, it was largely believed that the color vision system constantly recalibrates: The color vision system is trying to achieve color constancy on an ongoing basis. In order to do so, it needs to discount the illuminant, and to do that, it takes recent illumination history and even its optics and

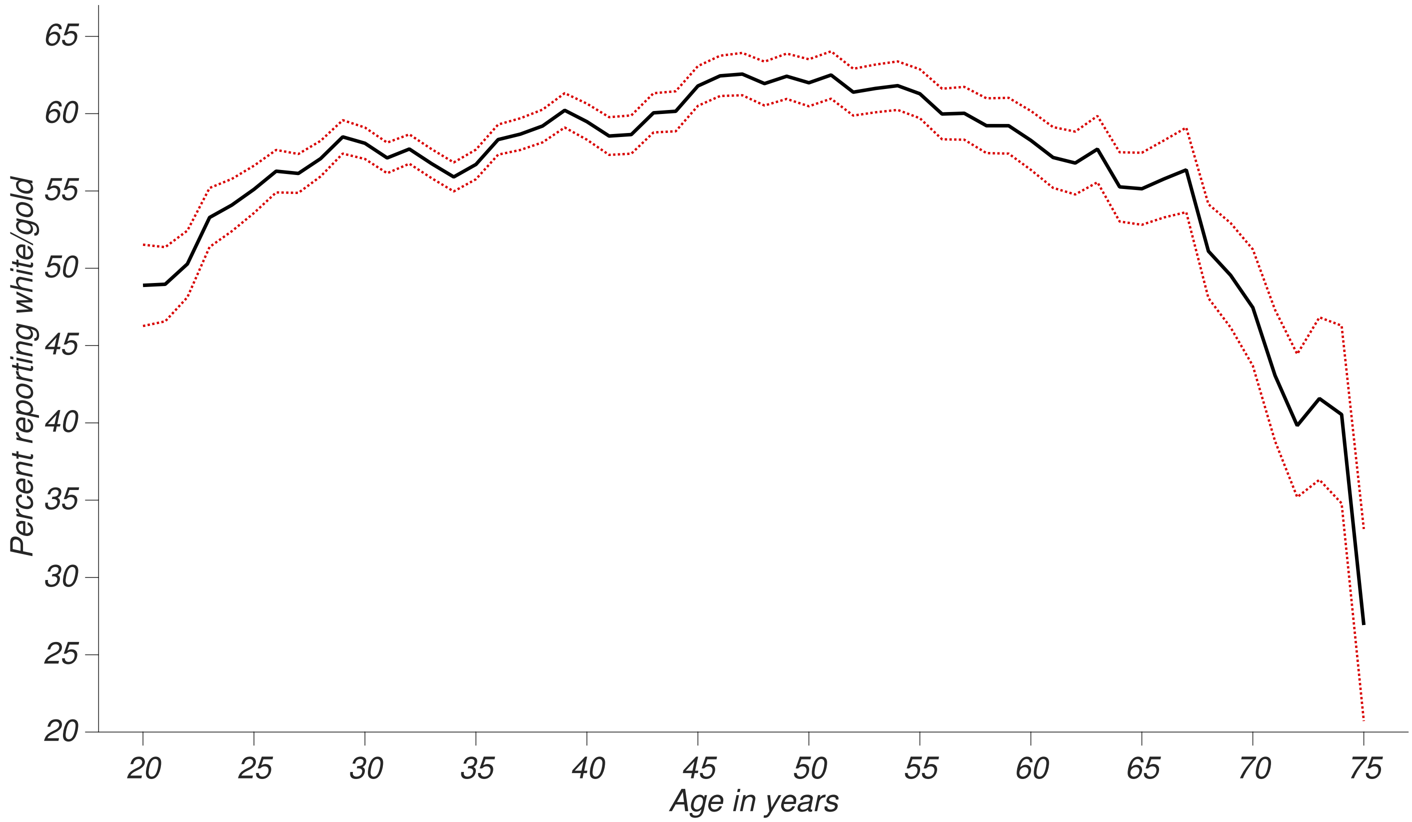

Figure 9. Color percept of the dress stimulus as a function of age from both runs. The *x*-axis represents the self-reported age of participants, blended with the two neighboring age bins. The *y*-axis represents the proportion of participants reporting seeing the dress stimulus as white and gold. The black line represents the mean response for a given age; the hashed red lines represent ± 1 *SEM*.

retina into account. Different retinal mosaics and optics don't seem to matter much across observers as suggested by color-matching experiments. As a consequence, observers—in the absence of profoundly abnormal color vision—are believed to more or less see colors in the same way even in the face of dramatically varying retinal mosaics, in particular differences in the ratio of L to M cones (Brainard et al., 2000). Indeed, this recalibration is believed to take place as a function of experience, but it usually takes rather drastic interventions to produce robust and lasting effects (Neitz, Carroll, Yamauchi, Neitz, & Williams, 2002). Our results suggest that differential color experiences might come about even as consequences of everyday living if lifestyle choices, such as when to rise and when to go to bed, create differential illumination priors in the population—an effect that might make itself felt, as in the case of the dress stimulus, when illumination is ill defined and color constancy mechanisms have to rely on priors to discount the—assumed—illuminant. Under those conditions, color experiences might be more idiosyncratic than previously assumed. In addition, illumination priors could account for the fact that for some people—a small minority—the dress is indeed bistable, rapidly flipping back and forth between different percepts. Maybe the illumination assumptions of these people are so precisely balanced that small perturbations—maybe even neural noise—can flip their percepts back and forth. It would be interesting to identify these individuals and study them further in this regard.

Whereas the field, including this paper, seems to converge on the interpretation that it is illumination assumptions that drive the dress effect (Witzel, Racey, & O'Regan, 2017), illumination priors might not be the only factor that determines the subjective color percept; there well might be other features that play a role, such as fabric priors. In the case of the dress itself, there are peculiar specularities that appear when the dress is illuminated by incandescent light, which could strongly influence the subjective percept of an observer (David Heeger, personal communication, January 30, 2016).

One way to verify whether it is really assumptions about illumination that matter the most would be to vary the illumination conditions of the dress image presented to observers in the lab. This might be as simple as comparing the illumination on the top of the dress image, implying natural light, with that on the bottom, implying artificial light, and seeing what kinds of color percepts of the dress this produces. In addition, one could compare these effects to those produced by specularities of different fabrics and illuminations. Of course one problem with all of this is that it will be increasingly hard to find naïve participants for this kind of research. Even in the original studies on the dress phenomenon, fewer than one in four participants was naïve to the dress image (Lafer-Sousa et al., 2015). By now, almost everyone will likely have seen the image of the dress, and if not, those potential participants are probably not representative of the population as a whole.

The fact that the field—despite best efforts—has thus far failed to come up with a stimulus that evokes comparable degrees of ambiguity suggests that we don't fully understand what drives the effect, and until we do so, we probably won't be able to probe what does because we will not be able to find enough new naïve participants.

Thus, it is a priority to come up with a stimulus display that produces similar degrees of ambiguity but is novel, perhaps by creating a situation with a deliberately ambiguous illuminant so that color constancy mechanisms will have to rely on—idiosyncratic—priors to discount the illuminant.

Conclusions

We believe we showed that differential assumptions about illumination account for a large part of the individual differences we see in subjective color experiences of the dress stimulus between observers. We think this is the case because, in all instances, querying beliefs about illumination modulated color perception strongly in accordance with theoretical predictions and—when appropriate—in a dose-dependent fashion with effects replicating across two data sets. In contrast, the effect of alternative predictors, such as demographic markers, like gender or age, were modest, hard to interpret, or both. In addition, our results are consistent with other mechanisms known to operate in the visual system, such as color constancy or the utilization of Bayesian priors.

Keywords: color vision, individual differences, Bayesian priors, dress, #thedress, power, survey, online, qualia, illumination, circadian type

Acknowledgments

We would like to thank David Heeger for discussion and help with getting internal review board approval to conduct this research. Parts of this study have previously been presented in abstract form at the annual meeting of the Society for Neuroscience 2015 in Chicago (Wallisch, 2015b) and the International Congress of Psychology in Yokohama, Japan, in 2016 (Wallisch, 2016).

Commercial relationships: none.
Corresponding author: Pascal Wallisch.
Email: pascal.wallisch@nyu.edu.
Address: Department of Psychology, New York University, New York, NY, USA.

References

Adan, A., Archer, S. N., Hidalgo, M. P., Di Milia, L., Natale, V., & Randler, C. (2012). Circadian typology: A comprehensive review. *Chronobiology International*, *29*(9), 1153–1175.

Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011, June 17). The visual impact of gossip. *Science*, *332*(6036), 1446–1448.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid A. M., Dale A. M., . . . Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences, USA*, *103*(2), 449–454.

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999, Dec 23). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, *402*(6764), 877–879.

Brainard, D. H., Roorda, A., Yamauchi, Y., Calderone, J. B., Metha, A., Neitz, M., . . . Jacobs, G. H. (2000). Functional consequences of the relative numbers of L and M cones. *Journal of the Optical Society of America A*, *17*(3), 607–614.

Brugger, P. (1999). One hundred years of an ambiguous figure: Happy birthday, duck/rabbit! *Perceptual and Motor Skills*, *89*(3), 973–977.

Changizi, M. A., & Hall, W. G. (2001). Thirst modulates a perception. *Perception*, *30*(12), 1489–1497.

Chetverikov, A., & Ivanchei, I. (2016). Seeing "the dress" in the right light: Perceived colors and inferred light sources. *Perception*, *45*(8), 910–930.

D'Zmura, M., & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A*, *3*(10), 1662–1672.

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Boca Raton, FL: CRC Press.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47.

Gegenfurtner, K. R., Bloj, M., & Toscani, M. (2015). The many colours of "the dress." *Current Biology*, *25*(13), R543–R544.

Gibertini, M., Graham, C., & Cook, M. R. (1999). Self-report of circadian type reflects the phase of the melatonin rhythm. *Biological Psychology*, *50*(1), 19–33.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, *9*(11), 1367–1368.

Hermann, L. (1870). Eine erscheinung simultanen contrastes [Translation: An appearance of simultaneous contrast]. *Pflügers Archiv Pflügers Archiv für die gesamte Physiologie 3*, 13–15.

Holderness, C. (2015). *What colors are this dress?* Retrieved from https://www.buzzfeed.com/catesish/help-am-i-going-insane-its-definitely-blue

Lafer-Sousa, R., Hermann, K. L., & Conway, B. R. (2015). Striking individual differences in color perception uncovered by "the dress" photograph. *Current Biology*, *25*(13), R545–R546.

Lotto, R. B., & Purves, D. (1999). The effects of color on brightness. *Nature Neuroscience*, *2*, 1010–1014.

Mahroo, O. A., Williams, K. M., Hossain, I. T., Yonova-Doing, E., Kozareva, D., Yusuf, A., & Hammond, C. J. (2017). Do twins share the same dress code? Quantifying relative genetic and environmental contributions to subjective perceptions of "the dress" in a classical twin study. *Journal of Vision*, *17*(1):29, 1–7, doi:10.1167/17.1.29. [PubMed] [Article]

McBride, B. (2011). *Labor Force Participation Rate by Age Group*. Retrieved from http://www.calculatedriskblog.com/2011/12/labor-force-participation-rate-by-age.html

Monnier, P., & Shevell, S. K. (2003). Large shifts in color appearance from patterned chromatic backgrounds. *Nature Neuroscience*, *6*(8), 801–802.

Mount, S. (2015). *No Simple Genetics for the Dress - But Does Color Vision Affect Compensation for Lighting?* Retrieved from http://ongenetics.blogspot.com/2015/03/no-simple-genetics-for-thedress-but.html

Necker, L. A. (1832). Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid. *The London and Edinburgh Philosophy Magazine and Journal of Science*, *3*, 329–337).

Neitz, J., Carroll, J., Yamauchi, Y., Neitz, M., & Williams, D. R. (2002). Color perception is mediated by a plastic neural mechanism that is adjustable in adults. *Neuron*, *35*(4), 783–792.

Ninio, J., & Stevens, K. A. (2000). Variations on the Hermann grid: An extinction illusion. *Perception*, *29*(10), 1209–1217.

Pegoraro, M., Picot, E., Hansen, C. N., Kyriacou, C. P., Rosato, E., & Tauber, E. (2015). Gene expression associated with early and late chronotypes in *Drosophila melanogaster*. *Frontiers in Neurology, 6*,100, http://doi.org/10.3389/fneur.2015.00100.

Rabin, J., Houser, B., Talbert, C., & Patel, R. (2016). Blue-black or white-gold? Early stage processing and the color of "the dress." *PloS One*, *11*(8), e0161090.

Rogers, A. (2015). *The science of why no one agrees on the color of this dress*. Retrieved from http://www.wired.com/2015/02/science-one-agrees-color-dress/

Salvi, S. M., Akhtar, S., & Currie, Z. (2006). Ageing changes in the eye. *Postgraduate Medical Journal*, *82*(971), 581–587.

Sathirapongsasuti, F. (2015). *Genetics and that striped dress*. Retrieved from https://blog.23andme.com/23andme-research/genetics-and-that-striped-dress/

Schlaffke, L., Golisch, A., Haag, L. M., Lenz, M., Heba, S., Lissek, S., . . . Tegenthoff, M. (2015). The brain's dress code: How the dress allows to decode the neuronal pathway of an optical illusion. *Cortex*, *73*, 271–275.

Smith, D. (2016). *Calculating the emission spectra from common light sources*. Retrieved from https://www.comsol.com/blogs/calculating-the-emission-spectra-from-common-light-sources/

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, *9*(4), 578–585.

Toscani, M., Gegenfurtner, K. R., & Doerschner, K. (2017). Differences in illumination estimation in #thedress. *Journal of Vision*, *17*(1):22, 1–14, doi:10.1167/17.1.22. [PubMed] [Article]

Vemuri, K., Bisla, K., Mulpuru, S., & Varadharajan, S. (2016). Do normal pupil diameter differences in the population underlie the color selection of #thedress? *Journal of the Optical Society of America A*, *33*(3), A137–A142.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*(4), 205–217.

Wallisch, P. (2015a). Brighter than the sun: Powerscape visualizations illustrate power needs in neuroscience and psychology. arXiv preprint arXiv: 1512.09368.

Wallisch, P. (2015b). Determinants of qualia of a profoundly ambiguous stimulus in the domain of color vision. Presented at the 45th Annual Meeting of the Society for Neuroscience, Chicago, IL.

Wallisch, P. (2016). Assumptions about illumination account for differential interpretations of "the dress" stimulus. Presented at the 31st International Congress of Psychology, Yokohama, Japan.

Wallisch, P., & Movshon, J. A. (2008). Structure and function come unglued in the visual cortex. *Neuron*, *60*(2), 195–197.

Watanabe, H., Fujii, T., Nakamura, T., & Korenaga, T. (2015, July). Color perception difference: White and gold, or black and blue? Presented as a poster at ACM SIGGRAPH 2015, August 9–13, 2015, Los Angeles, CA, http://dx.doi.org/10.1145/2787626.2787630.

Winkler, A. D., Spillmann, L., Werner, J. S., & Webster, M. A. (2015). Asymmetries in blue–yellow color perception and in the color of "the dress." *Current Biology*, *25*(13), R547–R548.

Witzel, C., Racey, C., & O'Regan, J. K. (2017). The most reasonable explanation of "the dress": Implicit assumptions about illumination. *Journal of Vision*, *17*(2):1, 1–19, doi:10.1167/17.2.1. [PubMed] [Article]

Yeshurun, Y., & Carrasco, M. (1998, Nov 5). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*(6706), 72–75.