Enabling Designers to Foresee Which Colors Users Cannot See

Katharina Reinecke*
Computer Science & Engineering, DUB Group
University of Washington
reinecke@cs.washington.edu

David R. Flatla*
School of Computing
University of Dundee
d.flatla@dundee.ac.uk

Christopher Brooks*
School of Information
University of Michigan
brooksch@umich.edu

ABSTRACT
Users frequently experience situations in which their ability to differentiate screen colors is affected by a diversity of situations, such as when bright sunlight causes glare, or when monitors are dimly lit. However, designers currently have no way of choosing colors that will be differentiable by users of various demographic backgrounds and abilities and in the wide range of situations where their designs may be viewed. Our goal is to provide designers with insight into the effect of real-world situational lighting conditions on people's ability to differentiate colors in applications and imagery. We therefore developed an online color differentiation test that includes a survey of situational lighting conditions, verified our test in a lab study, and deployed it in an online environment where we collected data from around 30,000 participants. We then created ColorCheck, an image-processing tool that shows designers the proportion of the population they include (or exclude) by their color choices.

Author Keywords
Color Vision; Situational Lighting Conditions; Color Differentiability; Design Software

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI)

INTRODUCTION
The inability to see on-screen content due to challenging lighting conditions is a commonly experienced but largely unexplored phenomenon. From the teenagers on the sidewalk contorting their bodies to cast shadows over their mobile phones to office workers adjusting their monitor settings for better readability, we are all aware of the impact that lighting, devices, and configuration settings can have on screen-based color differentiation. Previous work has shown that our color vision can be substantially distorted by a variety of *situational lighting conditions*, such as different combinations of bright sunlight, artificial lighting, glossy screens, and monitor settings [37] and that these factors can reduce perceived color contrast, inhibiting readability and information uptake [14].

However, designers have traditionally been largely unsupported when estimating the impact of users' varying color perception on their designs. This is aggravated by the fact that color vision is an inherently individual experience, influenced by both external factors, such as situational lighting conditions, and internal factors, such as acquired or inherited color vision deficiencies.

Current color design guidelines do not address this diversity in color perception, but are based either on aesthetics advice (e.g., Adobe Kuler [2], COLOURlovers [3]), or on design advice for people with color vision deficiencies (CVD) (e.g., [35, 1]). None of these guidelines incorporate the external factors that influence color differentiation ability, such as viewing a bright screen in a dimly lit room, or interacting with a glossy-screened device in bright sunlight.

To address this, we demonstrate how a population's color differentiation ability can be measured, modeled, and used by designers to understand the impact diverse situational lighting conditions have on color perception. In particular, we make the following contributions:

1. An open-sourced color differentiation test (WebCDT) that measures a user's color differentiation ability in a digital environment (e.g., on a computer, tablet, or phone). Through a controlled lab study, we demonstrate that WebCDT is sensitive to changes in environmental lighting.

2. An analysis of the variability in people's color vision based on around 30,000 participants who completed WebCDT under a variety of lighting conditions with various devices and monitor settings on the experiment platform LabintheWild.org. We evaluate the effect of situational and demographic conditions on participants' ability to discriminate colors, with major findings being that ambient and monitor brightness, age, gender, and self-reported color vision deficiencies all impact users' color differentiation abilities in digital environments. We also contribute a public dataset, which includes the color differentiation ability measurements of our 30,000 participants and their self-reported demographics, devices, monitor settings, and lighting conditions.

*All three authors equally contributed to this paper.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).
CHI'16, May 07–12, 2016, San Jose, CA, USA
ACM 978-1-4503-3362-7/16/05.
http://dx.doi.org/10.1145/2858036.2858077

3. A design tool, ColorCheck, that shows which parts of an image are likely not differentiable by certain proportions of users. We demonstrate how ColorCheck can be used to evaluate the effect of our participants' varying color differentiation abilities when viewing digital content using 450 website screenshots and 3,000 infographics.

RELATED WORK

The ability to differentiate colors is vital for information uptake in many areas, from reading maps and understanding information visualizations, to the interpretation of medical imagery. However, color differentiability can be severely impacted by a number of situational factors, such as monitor settings and lighting conditions [36] or perceptual abilities [8, 12]. Crucially, the effects of insufficiently differentiable colors can range from frustration to critical safety issues [8].

Effects of Situational Lighting Factors on Color Vision

Prior work has mentioned the importance of adjusting monitors and lighting in order to optimize viewing conditions. Ware [36, p.90], for example, suggests ensuring that "the room should have a standard light level and illuminant color" and that "only a minimal amount of light should be allowed to fall on the monitor screen" in order to perceive computer-generated colors similarly to colors in a room. This is of course difficult to achieve for users with handheld devices and laptops. Liu et al. [29] investigated the impact of ambient lighting on such handheld devices (tablets and mobile phones) used in the medical domain. They found that ambient lighting conditions (simulating dark, office, and outdoor environments) had a significant effect on visual task performance on mobile displays with participants performing best in dark conditions. Their findings support other research that has found the perceived image quality on mobile phone screens to decrease when ambient brightness levels were increased [23, 27, 28, 16]. Recently, Kim et al. [22] addressed the difficulties around knowing which colors online fashion products have when viewed in different lighting conditions by developing an approach to crowdsource color perception. By generating a "CrowdColor", their system approximates the "real" color by averaging users' color perception. The study used to evaluate CrowdColor was conducted in lab using two controlled lighting conditions and two different mobile devices; the current version of CrowdColor is therefore "limited to a controlled environment only" [22, p.483].

None of this prior work has explored users in their natural computing environments and in a variety of situational lighting conditions.

Effects of Demographics and Abilities on Color Vision

Color vision can be affected by a variety of inherited and acquired factors. Inherited Color Vision Deficiency (CVD) is a result of faulty or missing cones, the light-sensitive cells that facilitate color vision. Much work has studied inherited CVD and its impact on color perception (see, e.g., [8]). To mitigate the negative effects of CVD, Human-Computer Interaction (HCI) researchers have developed technology that addresses the lack of color differentiability, for example by recoloring user interfaces [15].

Unlike inherited factors influencing color vision, acquired factors are a result of environmental impacts or time. Damage to the brain, for example, can induce acquired CVD [10]. Age-related changes in human's perceptual system can lead to a declining ability to distinguish colors and a higher susceptibility to glare [12]. Our color vision can also be temporarily affected by certain prescription drugs [26].

A number of studies have also found that gender correlates with color vision [30, 4, 5]. For example, females have been found to better perceive changes in colors [4]. Females are also less likely to have inherited CVD: 1/12 men versus 1/200 women in the world are believed to have some kind of color vision deficiency [6] (although these numbers may be imprecise because many people are not aware of having CVD [8]).

While HCI researchers have mostly studied the influence of CVD on color perception in technology-mediated environments [25, 14, 15], our study extends prior work with the first large-scale analysis of people's varying color differentiation abilities linked to a combination of situational lighting conditions, age, gender, and CVD.

Color Differentiation Tests

To assess a person's color differentiation ability, a variety of tests typically aimed at detecting CVD have been developed. Paper-based color differentiation tests (such as the Ishihara plate test [20]), are used in clinical settings to identify whether a person has inherited CVD. These tests have been developed over the past one hundred years to become more specific in their diagnosis [11, 17], and were moved to computerized color vision tests in the 1990s [7, 32]. Each of these tests requires the user to distinguish between different-colored items (often dots).

Although technically sensitive to environmental fluctuations, most color vision tests assume that they are administered in an environment that controls for this potential source of noise, e.g., controlled intensity and type of ambient lighting for the paper-based tests, and on color-calibrated displays for the computerized tests. Recognizing the limitation of this assumption, Flatla and Gutwin [13] developed a new two-minute computerized color vision test designed specifically to gather computer users' real world color differentiation abilities. They showed that the test is sensitive to individual variations in color vision ability, such as induced by different monitor settings or lighting conditions [14]. As we are considering the influence of environmental factors on users' color differentiation abilities within uncontrolled environments (the Web), we based our online color vision differentiation test, WebCDT, on the most recent version of their test [14].

A WEB-BASED COLOR DIFFERENTIATION TEST

To develop a web-based test amenable to measuring a user's color differentiation ability in their current online environment, we adapted the color differentiation test by Flatla and Gutwin [14] in a number of ways. First, in their test, test takers are presented with a series of dynamic tessellating visual stimuli (four stills of these stimuli are shown in Figure 1), and are asked to identify the location of the gap in the circle. The test uses temporally-variant luminance noise (similar to

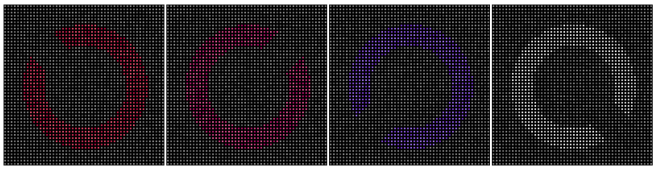


Figure 1: CDT stimuli: four stills of gapped circles at maximum difference from background in the computerized color differentiation test by Flatla and Gutwin.

'snow' on an unoccupied analogue television channel) by randomly re-assigning the background and circle color noise for each frame. Since this requires dynamically changing frames (i.e., videos) and would potentially exclude participants with limited bandwidth, we opted to replace the temporally-variant luminance noise with static pre-computed noise. Previous work on testing color vision online suggests that temporally-variant noise does not contribute to the validity of the test (and might be detrimental to some users), so can be safely removed [32]. We therefore programmatically captured a single randomly-chosen frame from the original test for each testing point along each hue and luminance axis.

Second, we halved the number of color axes tested by Flatla and Gutwin's test in order to keep the online test short. This also allowed us to be able to test the same color axis in multiple trials, which serves as an important verification in an uncontrolled online environment in that inconsistent color differentiation performance across the different trials would indicate that a participant's lighting conditions or monitor settings have changed while performing the test. In the original test, the intensity of the circle color is systematically changed to find the intensity level necessary to discriminate the circle from the background. The search is performed for six different hues (magenta, pink, yellow, green, cyan, blue) and two luminance axes (lighter and darker than the background gray) on the gray background shown in Figure 1. Because a perceptually-uniform color space is used to define the search colors, each chromatic axis is measured twice, in opposite directions out from gray (pink and cyan are in opposite directions, as are magenta and green, yellow and blue, and lighter and darker). In previous work, reducing the number of chromatic axes has been shown to have no significant impact on the predictive power of the resulting color differentiation models [14], thus we reduced the number of axes to four: pink, magenta, blue, and lighter, as these offer the largest ranges of colors to test.

The resulting WebCDT measures each of the four color axes three times, and takes about three minutes to complete. As in Flatla and Gutwin's test, the test identifies four differentiation limits that are used to construct a discrimination ellipsoid volume [14, 31] (short, *ellipsoid volume*) describing the color differentiation abilities of the user for the background gray color. Only colors outside of the ellipsoid volume are not differentiable from the background gray. Using a perceptually-uniform color space, the ellipsoid volume for the gray background can be adapted via affine transformations (translation and rotation) to be the ellipsoid volume for any other color, enabling predictions about the differentiability of any two colors (by finding the ellipsoid for one color and determining if the other color is inside or outside that ellipsoid). We use CIE L*u*v* color space to define our axis colors and ellipsoid volumes as this color space is perceptually-uniform (Euclidean distance maps to perceptual difference), and planes of continuous hue (radiating out from a central vertical achromatic axis) are flat within this color space (reducing the number and complexity of affine transformations required when performing color differentiation predictions) [34].

WebCDT Verification

A property of the ellipsoid volumes described above is that as the color differentiation abilities of a user get worse, his/her ellipsoid volume gets larger. To verify that the WebCDT is sensitive to variations in color vision despite the changes we made, we conducted a lab study in which we systematically adjusted the lighting conditions in which participants completed the WebCDT, and measured how this influences participant ellipsoid volumes.

Participants

A total of 14 volunteers were recruited from a local university (M=29.1 years, SD=10.1). Six were male, and half of our participants (three males) reported they had a pre-existing eye condition (e.g., myopia), that was corrected-to-normal.

Procedure and Experimental Design

Participants completed the WebCDT on a color-calibrated (Spyder4Express) laptop in a room with controlled lighting which remained on for the duration of the study. A lamp was positioned over the laptop, shining directly on the laptop screen. The lamp was illuminated for half of the trials, and turned off for the other half, giving two room lighting levels (1040.6 lux and 629.5 lux respectively). We measured the light levels at the laptop keyboard using a Sinometer Pocket Digital Light Meter LX1010BS.

For each trial, the laptop screen was set to one of its four darkest settings, corresponding to 0, 10%, 20%, and 30% brightness (as reported by the laptop display software). Using the light meter, we measured the brightness of the center of the screen with a pure white image showing to get the following light levels: 12, 24, 48, and 69 lux, respectively.

When using a screen (such as a laptop), the perceived brightness of the display is moderated by the brightness of the environment; a dark-screened laptop is fine in a dark room, but can be impossible to see outside on a sunny day. To take this into consideration, we calculated the ratio between the room lighting and the monitor brightness in lux. As lux is a perceptual measurement of brightness (i.e., it represents perceived intensity per unit area), this ratio is a unitless value that represents the difficulty of seeing items on the screen under those given lighting conditions (room and monitor brightness).

With four monitor brightness settings and two room lighting settings, we tested participants in the resulting eight different lighting ratios (9.1, 13.1, 15.1, 21.7, 26.2, 43.4, 52.5, 86.7) which served as our independent measures. Monitor brightness and lamp order were counterbalanced between participants. Participants performed the WebCDT in each of the

eight lighting ratios. We used the mean differentiation limit from the three repetitions of the WebCDT to generate the ellipsoid volume which served as our dependent measure.

As the lighting conditions grow increasingly more challenging (lighting ratio increases), the color differentiation abilities of participants should decrease (resulting in an increasing ellipsoid volume). As a result, we hypothesize that there should be a strong positive correlation between the lighting ratio and ellipsoid volume.

Results

We observed a significant relationship between lighting ratio (room lux/monitor lux) and participants' mean discrimination ellipsoid volumes (Pearson's $r = 0.97$, $p < .001$) as shown in Figure 2. Participants had a harder time discriminating between colors as the viewing conditions became more difficult. These results confirm our hypothesis that the ellipsoid volume increases proportionally to the lighting ratio. The results of our lab study therefore verify that the WebCDT is indeed sensitive to variations in environmental lighting conditions.

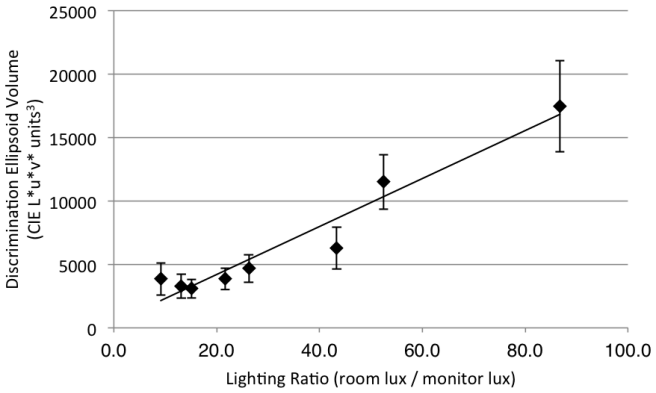


Figure 2: Mean discrimination ellipsoid volumes (in CIE L*u*v* units^3) ± s.e. for each of eight lighting ratios.

ONLINE EXPERIMENT

With WebCDT's sensitivity to changes in environmental lighting confirmed with in-lab participants, we proceeded to evaluate the color differentiation abilities of a more diverse population; diverse both in terms of the situational lighting conditions in their natural computing environments, and in terms of their demographics and perception abilities. More precisely, we wanted to answer two questions: (1) What is the variability in people's color differentiation abilities?, and (2) How do situational lighting conditions, demographics, and abilities impact their color differentiation abilities?

Method

To be able to recruit and study a large and diverse population, we deployed the WebCDT on the online experimentation platform LabintheWild (www.labinthewild.org). Participants came to the study via a link from the platform home page, other studies on the same site, and links posted by previous participants on social networking sites, blogs, or forums. Instead of receiving financial compensation, participants were able to test their "color age": Upon completion of the study, they were shown their estimated age based on a prediction model using their answers to the color vision test. This prediction served only as an incentive.

Procedure

After giving their informed consent, participants were asked to answer several questions about their situational lighting conditions as well as their computing setup. The type of lighting questions that we assessed were dependent in part on whether a participant indicated that they were indoors or outdoors. Since reporting these conditions can be subject to technical expertise, participants were given the option to indicate 'I don't know' for any of the questions.

Participants were then presented with instructions on how to perform the WebCDT, and were given the opportunity to perform as many practice trials as desired (practice trials were clearly labeled as such and excluded from analysis). As the study was only available in English but accessible to people around the world, we provided illustrations for most questionnaire items and the instructions of the test. The study can be accessed online at www.labinthewild.org/studies/color_age.

Each stimulus page included the WebCDT's stylized *C* image surrounded by 8 buttons to indicate the possible orientation of the image as shown in Figure 3. Participants were asked to click on one of these buttons to indicate the opening, or click "I can't tell" if they were unable to identify the position of the opening. Each of the four different color axes' were tested in a randomized order. This was repeated three times over (a 4x3 design) for a total of 12 color differentiation sets per participant. The number of *C* images shown per set varied dependent on participants' color differentiability. Participants were encouraged to take a break in the middle of the experiment.

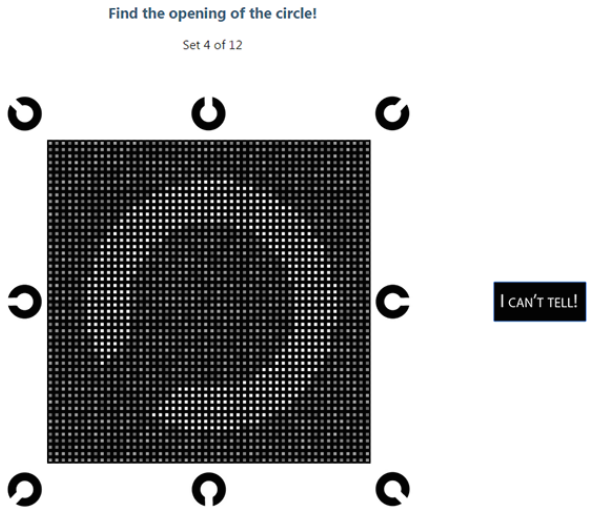


Figure 3: The main part of the experiment included a stimuli at the center of the page, a title to remind participants of the task, a subtitle indicating progress, as well as buttons for all eight possible options and "I can't tell".

The experiment concluded with a participant demographic questionnaire, including their country of residence, gender, age, highest education level, and whether they have any

known color vision deficiencies or corrected-to-normal vision. Participants were also asked to reveal whether this was the first time that they were taking this test. All information except for the retake question were voluntary. Participants were also given the opportunity to report on any technical difficulties, lighting changes, or other distractions that may have influenced their performance in the experiments. The entire study took 5-8 minutes to complete.

Participants

The experiment was deployed online between July 17, 2014 and July 28, 2015 and completed 31,248 times. The data of 1,765 participants was removed due to self-reported technical difficulties (e.g., Internet connection issues), changes in lighting conditions, or distractions. We excluded an additional 439 participants who reported to have participated in the experiment before. The following analysis therefore reports on the data of 29,044 participants.

Participants were between 5-94 years old (M=30.2, SD=15.2) and 70.6% were female. The plurality of participants (40.8%) reported to be enrolled or completed college, 22.8% currently attended high school, and 18.6% graduate school. The remaining participants were enrolled in professional schools (7.6%), pre-high school (3.8%), or had finished a PhD education (6.4%). Participants reported to come from 187 countries. Two percent of male participants reported to have some kind of CVD, as opposed to 1.7% of females.

Results

We first evaluated the variety of settings and ambient lighting conditions that participants were exposed to when taking the experiment, before analyzing the influence of these factors on their color differentiation abilities.

Range of Devices and Monitor Settings

Asked about the type of monitor, most participants reported that they were using built-in laptop monitors (47.33%), 26.48% used PCs with external monitors, and 26.24% were using monitors integrated in tablet computers or mobile phones. This suggests that at least 73.6% of participants were using monitors adjustable in tilt-angle, which could result in varying color vision abilities even throughout the test. Participants additionally revealed that their screens were mostly matte (63.8%) as opposed to glossy (36.2%).

We also asked participants to report on the brightness of their display, given a 0 to 100 point continuous scale from "Minimum brightness" to "Maximum brightness". The mean monitor brightness was 67.49 (SD=92.51). A positive correlation between monitor brightness and ambient brightness (Pearson's $r_{(15072)} = .18$, $p < .0001$) suggests that people set their monitors to higher brightness levels in bright lighting conditions, and could also reflect automatic brightness adaption on mobiles and tablets.

Breadth of Situational Lighting Conditions

The majority of participants (76.6%) indicated that they conducted the experiment indoors. To assess their situational lighting conditions, we asked participants "How bright are your surroundings?" and provided a continuous scale from "Midnight pitch black" to "Noon summer sun" which was mapped to values 0 through 100 respectively. Participants reported a significantly higher ambient brightness if they were outdoors (M=49.08, SD=22.75) than if they were indoors (M=46.54, SD=24.05, independent two-tailed t-test: $t_{(14799)} = -5.45$, $p < .0001$)[1].

In addition, participants were able to indicate their ambient lighting color temperature on a continuous scale from "Yellow" to "White blue", again mapped to values 0 through 100. The average ambient color value was 38.68 (SD=25.25), suggesting that the distribution of ambient color values was slightly skewed towards a more yellow tone. The ambient color reported by outdoor participants (M=37.42, SD=24.0) was significantly lower from the average value indicated by indoor participants (M=38.48, SD=25.07, independent two-tailed t-test: $t_{(13381)} = 2.09$, $p < .05$)[2].

Variability in Participants' Color Differentiation Abilities

To assess the variability in our population's ability to differentiate colors, we first computed participants' ellipsoid volumes based on their responses in the WebCDT. Figure 4 shows a histogram of participants' ellipsoid volumes with a mean ellipsoid volume of 3670.43 (SD=13728.03). The distribution of these ellipsoid volumes is positively skewed; 75% of participants have ellipsoid volumes below 3223.60. This is consistent with the results that our in-lab participants achieved when tested in conditions with lower lighting ratios (i.e., when the ratio between the room lighting and the monitor brightness was 9.1, 13.1, or 15.1, which corresponds to easier viewing conditions). We can therefore assume that the majority of our participants conducted the test under reasonable viewing conditions in which they were able to differentiate colors well. Nevertheless, a long tail in the distribution suggests that 25% of participants had poor viewing conditions and/or other factors influencing their color vision (factors that could explain this variance are analyzed in the next section).

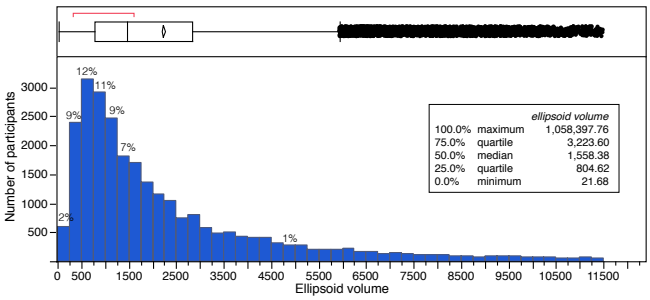


Figure 4: Histogram of number of participants versus mean discrimination ellipsoid volumes. Participants with ellipsoid volumes above 12,000 are not shown for space reasons.

In general, the larger an ellipsoid volume for a person, the fewer colors that person is able to differentiate. However, the orientation and shape of an ellipsoid volume within the CIE L*u*v* color space will influence the number of differentiable sRGB colors, because the sRGB color gamut is not

[1] CI outdoors [48.31-49.85], indoors [46.10-46.98]
[2] CI outdoors [36.59-38.26], indoors [37.99-38.97]

regular within L*u*v* [34]. With this in mind, we estimated the number of differentiable colors for a variety of ellipsoid volume orientations, shapes, and sizes by packing the ellipsoids inside the sRGB gamut within the L*u*v* color space, thereby generating upper and lower estimates of the number of differentiable colors for a given ellipsoid volume (Figure 5). Note that color differentiation is not linearly related to ellipsoid volume, and that the number of distinguishable colors drops significantly as we approach the mean ellipsoid volume (first vertical line) of our participant pool.

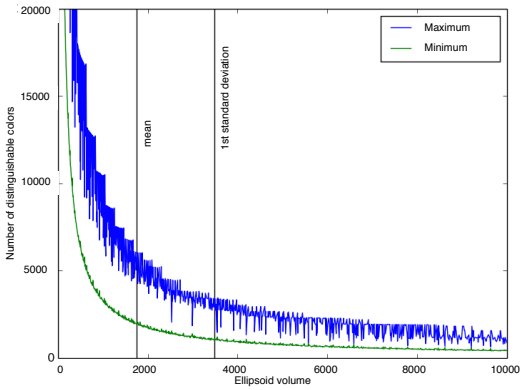


Figure 5: Estimate of the number of unique differentiable colors versus discrimination ellipsoid volume. Minimum and maximum values for color gamut based on ellipsoid shape are shown as green and blue lines respectively. Vertical lines show positions of the mean and the first standard deviation for our participant population.

Relationship with Situational Lighting and Demographics

In order to reduce error from participants' changing situational factors during the test (e.g., turning on lights, changing monitor settings) we discarded data from participants whose performance was highly variable in the WebCDT within the three runs for each vector (i.e., if one or more vectors resulted in one standard deviation outside of the mean of all participants). The data of these participants would be misleading to use in this analysis because we are relating their lighting conditions (self-reported at the beginning of the test) to their responses throughout the test. Note that this variability in responses within a vector does not occur in controlled lab settings where situational lighting conditions are held constant. As a result, 5,924 participants (20.4%) were excluded; the following analysis therefore reports on 23,120 participants with steady color vision responses across the three trials for each vector. The reduced dataset has a mean ellipsoid volume of 2138.71 (*SD*=2927.36), so slightly less variability in ellipsoid volumes than was the case for our entire population. We note that the mode is the same as before (453.72), indicating that while the long-tail was removed, the most frequently occurring ellipsoid volume level remained constant.[3]

To analyze this data, we fitted a multiple linear regression model using the ellipsoid volume as the dependent variable. We included demographic variables if there was a theoretical basis that suggests a possible impact on people's color vision, as it is the case for age [24], gender [5], and color vision deficiencies [6]. Lighting conditions, devices, and demographic information were modeled as independent variables.

The results show that age ($F_{(1)}$=61.06, $p < .0001$), gender ($F_{(1)}$=17.31, $p < .0001$), and whether or not a participant reported to have a color vision deficiency ($F_{(1)}$=164.19, $p < .0001$) have significant main effects on participants' ellipsoid volumes (see Table 1 for parameter estimates of the regression results). Ambient brightness ($F_{(1)}$=19.85, $p < .0001$) and monitor brightness ($F_{(1)}$=10.82, $p < .001$) also showed significant main effects. Despite the significantly larger average ellipsoid volumes for outdoor participants (*M*=2205.48, *SD*=3273.86) than for those participants who accessed the experiment indoors (*M*=2109.07, *SD*=2817.66, two-tailed t-test $t_{(20418)} = -1.99, p < .05$)[4], whether or not someone was indoors did not significantly contribute to the model fit. This is most likely because the variance in ellipsoid volumes between indoor and outdoor participants is already accounted for by participants' ambient brightness levels in the model. Participants' screen surface (matte or glossy) and ambient color also did not have significant main effects.

Table 1: Regression model analyzing the influence of situational lighting conditions, devices, and demographics on participants' ellipsoid volumes, $R^2 = .06, p < .0001, F_{(8,4077)} = 33.15, p < .0001$.

	β	SE β	t Ratio	$p <$
Constant	2532.05	187.25	13.52	.0001
Age	21.98	2.81	7.81	.0001
Gender [male]	-167.37	40.23	-4.16	.0001
CVD [none]	-1530.38	119.32	-12.81	.0001
Setting [outdoors]	1.78	43.42	0.04	n.s.
Glossy screen [false]	35.60	37.79	0.94	n.s.
Ambient color	0.61	1.44	0.42	n.s.
Ambient brightness	7.02	1.58	4.45	.0001
Monitor brightness	-4.10	1.25	-3.29	.001

The slight but significant positive correlation between age and ellipsoid volume that we found in our data ($\beta = 21.98, p < .0001$) confirms previous findings in the literature [24]. We found that with every year in age, the ellipsoid volume increases by 22.72 units with participants' color differentiability being best between 20 and 30. With a mean ellipsoid volume of 2138.71, this is, on average, a mostly negligible increase.

In line with previous work [5], we also found that males have larger ellipsoid volumes than females (β =-167.37, $p < .0001$), despite the fact that the model controls for the slightly higher occurrence of CVD in males. Participants who reported having inherited CVD have larger ellipsoid volumes than those without (β=-1530.38, $p < .0001$).

Finally, our results show that ambient brightness is positively correlated with ellipsoid volume ($\beta = 7.02, p < .0001$). Color differentiation ability gets worse as the surrounding light gets

[3]Note also that we have no way of differentiating whether a high within-vector variability is due to changes in lighting conditions or distractions (e.g., a participant not fully focusing on the test).

[4]CI outdoors [2122.3-2298.7], indoors [2065.0-2153.2]

brighter. Similarly, monitor brightness is negatively correlated with ellipsoid volume ($\beta = -4.10$, $p < .001$); setting the monitor to a low brightness has a slight negative impact on people's color differentiation ability.

While both of these findings were expected, the magnitude of these results (7.02 and -4.10 with a mean ellipsoid volume of 2138.71) was surprising. There are two possible reasons why situational lighting conditions do not play a larger role in contributing to the model fit. First, our participants might have avoided extreme situational lighting conditions when participating in our color differentiation test (self-selection). Thus, even though we see a higher variation in participants' color differentiation ability if they reported to be in bright lighting conditions versus darker surroundings, this variation is likely to be *higher* in actual day-to-day interactions when they have less freedom to opt-out of color differentiation tasks (e.g., when working with older displays outside). Second, we asked participants to self-report their situational lighting conditions, which is most likely subject to individual perceptions, or may not have fully captured people's ambient light. It is likely that the influence of situational lighting conditions is larger than our data shows due to the fact that much of the variation in participants' color differentiation ability is currently unaccounted for (the current model only explains 6% of the variance in participants' ellipsoid volumes).

In summary, our analysis showed a large variability in color differentiation abilities of a broad web population and showed that participants' age, gender, color vision deficiencies, as well as ambient brightness and monitor brightness significantly influence color differentiability. In the next section, we will use this data to demonstrate how designers might improve their practice around color use.

COLOR DIFFERENTIATION IN PRACTICE

While our previous results indicate a wide range in people's color differentiation abilities, we have yet to evaluate what this means in daily life. What are the consequences of this large variability in color differentiation abilities? Which part of the population is potentially excluded from properly viewing digital content as a designer might have intended?

To address these questions, we developed ColorCheck, a tool that evaluates an image and provides a visual cue to the designer (in the form of an image mask) as to the limitations of color differentiation. We then applied ColorCheck to two image datasets (websites and infographics) to characterize how the color differentiation ability of our web population affects viewing the colors in these images.

ColorCheck: A Tool to Show Color Differentiability

ColorCheck is an open-source image processing tool[5], developed to support design decisions. It employs an algorithm that compares pairs of colors within a color space and determines if one color occupies the ellipsoid volume of the other color. Pairs of colors which overlap in this manner are deemed to be non-differentiable. There are often areas of an image that contain subtle fluctuations in color that ColorCheck may potentially identify as 'problem colors' when in reality, being able to precisely differentiate these colors is not important (e.g., the fluctuations in lighting of an object in an outdoor scene). To help reduce the effects of these subtle variations, ColorCheck discretizes the colors of each input image. To accomplish this, it identifies the sRGB gamut within the CIE L*a*b* color space, and creates bins within this space each with a radius of 5 L*a*b* units (giving 826 bins). Each sRGB color is then mapped to its most perceptually-similar bin and represented using the sRGB color at the center of this bin. We used a D65 illuminant when converting from sRGB to L*a*b*, and employed Euclidean distance to represent perceptual similarity, as the comparison distance is small compared to the total gamut of L*a*b* color space. In addition to reducing the effects of subtle fluctuations in color, discretization also improves the computational efficiency of ColorCheck, allowing individual images to be processed on demand (typically < 350 ms per image on a 2.8 GHz late-2013 MacBook Pro).

(a) Original image

(b) Color pairs that are not differentiable by 15% of the population are masked.

Figure 6: ColorCheck creates masked images, which black out color pairs that are not differentiable by a certain percentage of the population.

By using a color differentiation model collected from a broad population (such as the ellipsoid volumes computed from our cleaned dataset of 23,120 participants), ColorCheck is able to measure the proportion of the population that is unable to differentiate any pair of image colors. ColorCheck identifies the ratio of pixels within the input image that are not differentiable. The output of this process is a data file containing a two dimensional description of color differentiability for a given proportion of the population (e.g., one value in this description might be that 60% of a given image's colors cannot be distinguished by 20% of the population).

In addition to this broad understanding about how an image may be seen by a population, ColorCheck is able to layer a mask (shown as black pixels) on top of an image which shows the color differentiability for a specific part of the population. Using this, designers can easily identify the areas of an image that contain problematic color pairs. This is demonstrated in Figure 6: ColorCheck takes an original image (Figure 6a) and blacks out those pixels that are not differentiable by a given percentage of the population (Figure 6b).

Note that these masked images are not intended to show how an image would be perceived by the population – this would be a difficult task given the variability of color perception

[5]ColorCheck can be downloaded at www.labinthewild.org/data

within the population. Instead, ColorCheck allows designers to see if the critical parts of their images are perceivable (in terms of color differentiability) by their target population: Color-coded regions that are blacked out usually suggest that users lose important visual cues to understand data. If adjacent colors are black, users will be unable to separate content regions. Finally, if large parts of an image are set to black, designers can assume that users will be unable to decipher the website or image's message, and aesthetic appreciation of the image could be significantly altered.

In practice, we expect that designers would view the same image with various differentiability thresholds, for example that of 50% or 80%. In the images generated by ColorCheck, all color pairs that are not differentiable by 50% and 20% of the population (respectively) will have been set to black. A designer can then use this information to decide whether the masked parts are essential for interpreting the content. If yes, they may decide to recolor critical areas, such as by choosing colors that are further apart in the color space.

In its current form, ColorCheck leaves the decision which color pairs need to be differentiable to the designer. This decision was made because of the difficulty to computationally determine the intentions of designers, and in particular, which colors designers necessarily require to be differentiable. To support designers in this process, ColorCheck provides a batch mode for processing 0-100% of the population, generating a separate masked image for each percentage. By visually scanning each image in increasing percentage order, pairs of problem colors can be identified when both regions of the image become masked at the same percentage level.

As ColorCheck relies on a data set of ellipsoid volumes, the output currently shows which parts of an image are not differentiable by our web-based participant population. This sample might not be representative of the average computer user, and likely does not balance usage among situational lighting conditions and devices. Our experiment therefore remains online; we plan to regularly update the tool with the resulting data to achieve more representative forecasts of users' ability to differentiate colors. In addition, ColorCheck users can plug in their own data sets (e.g., the ellipsoid volumes of their target group obtained through the use of WebCDT). It is also possible to reduce the input data set to a specific population of interest based on demographics, device and monitor settings, situational lighting conditions, or color vision deficiency (e.g., users over a certain age or those on mobile devices). Designers can sub-sample our open-access dataset, or use WebCDT to generate their own custom dataset.

Color Differentiability in Websites and Infographics

We employed ColorCheck to evaluate the level of differentiation between two example image datasets: (1) 450 website screenshots, selected to represent a range of domains and to vary in colorfulness, and (2) 3,000 infographics from the online community visual.ly.[6]

[6]These datasets have been previously published in [33] and [18] and were used with permission.

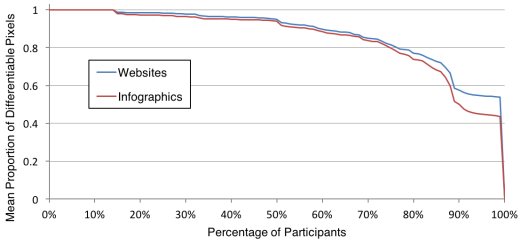


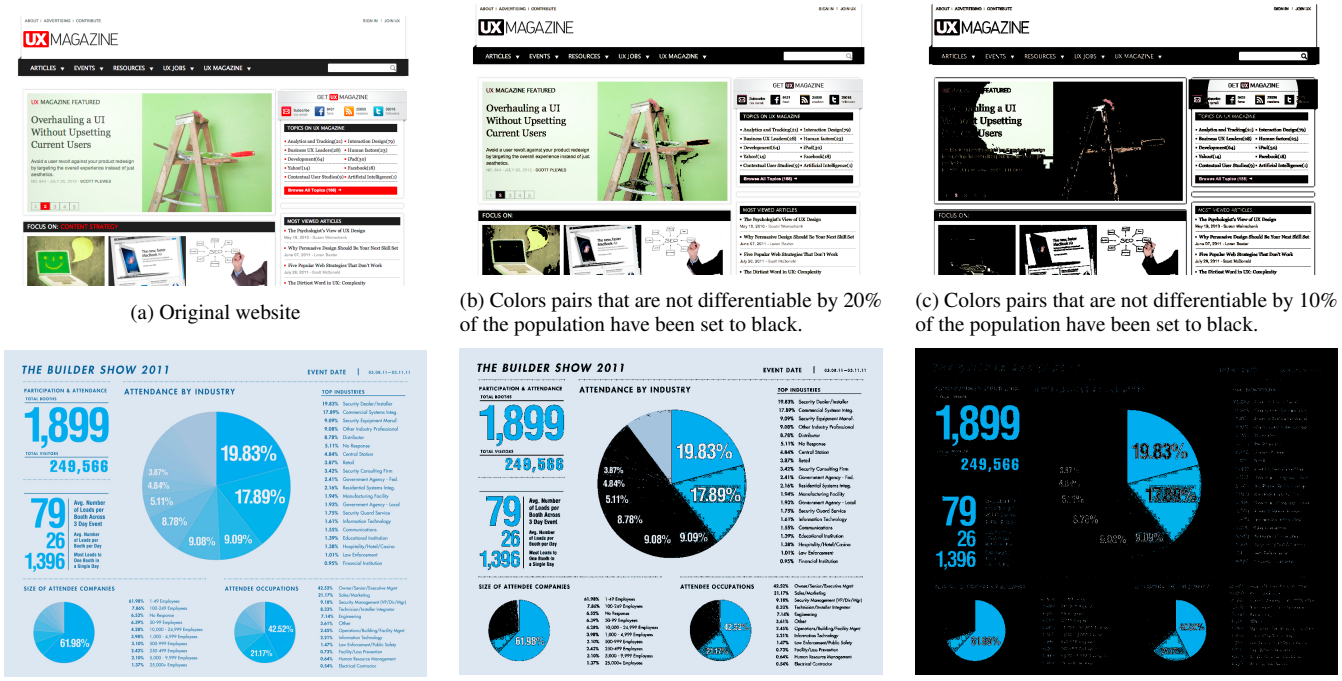
Figure 7: Mean proportion of image pixels differentiable for different percentages of the population.

To estimate how many images in our datasets contain colors that are non-differentiable for different percentages of our participant population, we calculated the average proportion of differentiable colors within each image dataset (across all websites, and across all infographics, respectively). The results of this analysis shows that 12% of our population can differentiate all colors (100% of pixels) in the websites and infographics in our dataset (see Figure 7). These are participants whose ellipsoid volumes were below 500 (see also Figure 4 for the distribution of ellipsoid volumes). Ellipsoid volumes larger than 500 result in a steady loss of differentiability in our image datasets. Roughly half of the population (52%) is unable to differentiate 10% of the colors in an average website or infographic. As we increase the proportion of the population we are targeting, the mean differentiability further decreases to a near-plateau between 90% and 99% of the population differentiating just under 60% of website content and just over 40% of infographic content. One of the reasons for infographics to perform more poorly than websites is that infographics often use homogeneous color schemes, so neighboring colors vary only in intensity. We also observe a sharp decline when targeting more than 99% of the population, which are people with severe color differentiation difficulties in our dataset (i.e., those whose test data resulted in extremely high ellipsoid volume levels).

Table 2: Comparison of situational lighting conditions and demographics of the 10% participants with the worst color differentiability, and the remaining 90%. Tests of the equality of two proportions report on Pearson's chi-squared test.

Factor	Worst 10%	Top 90%	Statistics
Outdoors	30.05%	23.14%	$\chi^2_{(1,20420)} = 5.38, p < .05$
Monitor brightness	M=62.81, SD=27.96	M=68.22, SD=29.10	$t_{(106.03)} = -1.97, p < .05$
Ambient brightness	M=53.81, SD=22.84	M=47.10, SD=23.79	$t_{(142.93)} = 3.50, p < .001$
Age	M=40.29, SD=20.10	M=30.21, SD=14.83	$t_{(111.04)} = 5.27, p < .0001$
Male	37.39%	26.58%	$\chi^2_{(2,23120)} = 13.64, p < .01$
CVD	23.04%	2.24%	$\chi^2_{(1,23120)} = 413.50, p < .0001$

But who are the 10% of participants that are most severely affected by a lack of color differentiability in websites and infographics? From our earlier analysis, we know that participants who reported being in brighter surroundings, and those who had set their monitors to lower brightness values performed worse. We also saw that being older, male, and



(a) Original website

(b) Colors pairs that are not differentiable by 20% of the population have been set to black.

(c) Colors pairs that are not differentiable by 10% of the population have been set to black.

(d) Original infographic

(e) Colors pairs that are not differentiable by 20% of the population have been set to black.

(f) Colors pairs that are not differentiable by 10% of the population have been set to black.

Figure 8: Image masks generated by ColorCheck for websites and infographics to show which colors become indistinguishable for certain proportions of the population.

having CVD has a negative effect on color differentiability. Those 10% of participants who are most severely affected by not being able to differentiate roughly 50% of the pixels in an average website and infographic confirm these characteristics; they are significantly more likely to have taken the test outdoors compared to the remaining 90% of participants (see Table 2 for statistics), they reported to have a lower average monitor brightness, as well as higher ambient brightness levels. In addition, the most affected 10% are, on average, ten years older than other participants, more likely to be male, and much more likely to have a color vision deficiency.

Figure 8 shows the actual effect of this loss of color vision with the original images (website and infographic) as well as their masked counterparts with non-differentiable colors for 20% and 10% of the population, respectively. For 10% of our participants, both the example website and infographic become almost entirely illegible. Information uptake would be severely impacted, and even a basic understanding of the content might be impossible. Setting the threshold to 80% (i.e., masking colors that are not differentiable by 20% of the population) improves this; in the case of the example website shown in Figure 8b, 20% of our participants will not be able to differentiate the colors of the social media buttons on the right, as well as several color accents used by the designer. While the general content remains viewable, such lack of color differentiability could have critical consequences. First, colors are often used as visual cues to guide the user's attention to specific parts of a platform. In the case of this website, roughly 20% of users might miss the opportunity to "browse all topics", because the button does not stand out from other headlines. Second, colors contribute to branding and an overall perceived appeal of an image (see, e.g., [9, 33]). If colors, such as the red used in the "UX Magazine" logo, become indistinguishable from others on the same website, users are presented with a site that essentially uses a reduced color scheme, resulting in a lower overall colorfulness. As prior work has shown, such changes in the level of colorfulness can impact users' perceived appeal [33].

The situation is especially severe in cases where reduced color differentiation ability impacts decoding information, such as when viewing data visualizations in difficult lighting conditions. The infographic in Figure 8d, for example, relies on viewers being able to distinguish different shades of blue in a pie chart. However, as Figure 8e shows, this is almost entirely impossible for 20% of our participants. For 10% of them (see Figure 8f), the dark blue text on a light blue background even becomes entirely illegible.

These examples are representative of many other websites and infographics in our dataset, suggesting that large proportions of users have to accept a compromised user experience in their daily interactions with digital content (see www.labinthewild.org/data for more examples).

LIMITATIONS

The main results of our study are based on a sample of approximately 30,000 online participants, who may not be representative of general computer users. We hope that leaving the WebCDT experiment online to enable frequent updating

of the ellipsoid volumes used in ColorCheck will result in an increasingly representative population, and a better representation of the color differentiation abilities for any user.

Our results are also limited by the fact that our websites and infographics datasets might not be representative of an average website/infographic. Thus, any conclusions on the percentage of colors our participants are unable to differentiate on average have to be seen in the context of these datasets.

SUMMARY AND DISCUSSION

While color choices have been the subject of much prior work, it has previously been impossible to foresee the consequences of these choices in increasingly diverse digital environments. The goal of this work was to quantify how many users cannot differentiate colors in a given user interface or image (e.g., in websites or infographics) and varying situational lighting conditions. To do so, we developed WebCDT, an online color differentiation test, verified that it is sensitive to changes in situational lighting conditions, and deployed it online. The results of a heterogeneous sample of around 30,000 online participants showed that (1) their color differentiation abilities vary considerably, and (2) situational lighting conditions (in particular, being outdoors, and/or having a low monitor brightness and high ambient brightness levels), demographics (age and gender) and possible color vision deficiencies have significant effects on color perception.

To evaluate the effect of our population's wide range of color differentiation abilities on viewing colors in digital environments, we developed an image-processing tool that can also be used by designers. ColorCheck identifies pairs of colors in an image that are not differentiable by certain proportions of a population. Applying ColorCheck to commonly viewed digital content (websites and infographics), we showed that 88% of our participants are unable to differentiate some colors in websites and infographics. Crucially, 10% of our participants can only distinguish around 50% of the colors in an average website/infographic.

We observed that the proportion of users that are affected by a reduced color differentiation ability is *larger* than what was previously assumed; it is not only people with (color) vision impairments, but any user who accesses digital content in suboptimal lighting conditions. As a result, large numbers of users could experience usability issues (e.g., if users cannot perceive color-coded cues in an interface), obstacles in information uptake (e.g., if color-coding in charts hinders data interpretation), or a reduction of perceived appeal (e.g., if an image is perceived with a reduced colorfulness). In practice, all of these issues could have serious consequences for both the user and the designer; for users, an inability to differentiate colors can have consequences for safety, learning and information uptake, or simply be frustrating [8]. For designers, possible consequences could be that people might switch to competitors, misunderstand their message, or simply will not appreciate their designs.

Ideally, we would therefore provide designers with concrete guidelines on how to design for the whole spectrum of people's color perception. From our analysis of infographics and websites, we have seen that homogeneous color schemes did especially poorly, because users will not be able to differentiate neighboring colors. In comparison, starkly contrasting colors (i.e., those that are far apart in most color spaces) will fare better, but using them is not always feasible if certain color schemes are favored by branding and/or aesthetic guidelines. There is an inevitable trade-off between aesthetics and designing for a broad range of (situational) color differentiation abilities, which necessarily excludes the option of prescriptive guidelines. Hence, ColorCheck was only developed as a demonstration tool, one that visualizes difficult-to-distinguish parts of an image. The decision of how to recolor (parts of) the image is deliberately left to the designer, and has to be made with three thoughts in mind: (1) Who is my target audience, and in which situations do I expect them to view digital content?, (2) Can I freely choose colors or do I have to adhere to specific color schemes and branding guidelines?, and (3) How important is it that users are able to differentiate the problem colors identified by ColorCheck?

While ColorCheck provides initial support for identifying problem color pairs, a future extension could include the ability to interact with the image masks (e.g., by mousing over a certain color) to more easily identify which other parts of the image are not differentiable. We also plan to integrate population characteristic filtering (e.g., to check the effect on a specific age group). Finally, we envision a mode that enables auto-generation of possible replacement colors, such as based on current viewing conditions. With the help of such automated recoloring of images, users with optimal color vision could still appreciate the aesthetics of the image as the designer intended, but others could be supported with colors that are further apart in the color space. Technically, calculating glare with the help of a tilt-sensor, such as employed in [19], in addition to information about time of the day and sunlight/cloudy skies, could aid recoloring decisions. Enlarging icons and adjusting the background color has previously been shown to improve readability for users with vision impairments [21] and could therefore be one of the possible image adaptations that is provided to users. We are excited to explore such automatic adaptations as part of our future work.

In summary, we have demonstrated that color choices can limit the ability of large proportions of users to perceive images as the designer intended. This work emphasizes that an inability to differentiate colors does not only stem from inherited or acquired CVDs—as commonly explored in HCI—but that situational lighting conditions can make *anybody* temporarily 'colorblind'. With laptops and mobile devices becoming increasingly ubiquitous, we believe that it is time digital content is designed accordingly.

DATASET AND SOURCE CODE

Our participant dataset, as well as the source code for WebCDT and ColorCheck can be accessed at www.labinthewild.org/data.

ACKNOWLEDGMENTS

We thank our in-lab and LabintheWild participants, Tiffany Liu for helping with the LITW implementation, and Daniel Herron for running the WebCDT verification study.

REFERENCES

1. 2008. Web Content Accessibility Guidelines. http://www.w3.org/TR/WCAG20/. (2008). Accessed: 2015-09-10.
2. 2015. Adobe Kuler. http://color.adobe.com. (2015). Accessed: 2015-09-10.
3. 2015. COLOURlovers. http://www.colourlovers.com. (2015). Accessed: 2015-09-10.
4. Israel Abramov, James Gordon, Olga Feldman, and Alla Chavarga. 2012. Sex and vision II: color appearance of monochromatic lights. *Biology of Sex Differences* 3, 1 (2012), 21.
5. David L Bimler, John Kirkland, Kimberly A Jameson, and others. 2004. Quantifying variations in personal color spaces: Are there sex differences in color vision? *COLOR Research and application* 29, 2 (2004), 128–133.
6. Jennifer Birch. 2001. *Diagnosis of Defective Colour Vision* (second ed.). Butterworth Heinemann, Linacre House, Jordan Hill, Oxford.
7. Jennifer Birch, John L. Barbur, and Alister J. Harlow. 1992. New Method Based on Random Luminance Masking for Measuring Isochromatic Zones Using High Resolution Colour Displays. *Ophthalmic & Physiological Optics: the Journal of the British College of Ophthalmic Opticians (Optometrists)* 12, 2 (April 1992), 133–136.
8. Barry L Cole. 2004. The handicap of abnormal colour vision. *Clinical and Experimental Optometry* 87, 4-5 (2004), 258–275.
9. Dianne Cyr, Milena Head, and Hector Larios. 2010. Colour appeal in website design within and across cultures: A multi-method evaluation. *International Journal of Human-Computer Studies* 68, 1 (2010), 1–21.
10. Finlay Dick, Sean Semple, Ruoling Chen, and Anthony Seaton. 2000. Neurological deficits in solvent-exposed painters: a syndrome including impaired colour vision, cognitive defects, tremor and loss of vibration sensation. *QJM* 93, 10 (2000), 655–661.
11. Dean Farnsworth. 1943. Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision. *Journal of the Optical Society of America (1917-1983)* 33 (October 1943).
12. Arthur D Fisk, Wendy A Rogers, Neil Charness, Sara J Czaja, and Joseph Sharit. 2009. *Designing for older adults: Principles and creative human factors approaches*. CRC press.
13. David R. Flatla and Carl Gutwin. 2011. Improving calibration time and accuracy for situation-specific models of color differentiation. In *ASSETS '11: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*. 195–202.
14. David R. Flatla and Carl Gutwin. 2012. SSMRecolor: Improving Recoloring Tools with Situation-specific Models of Color Differentiation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. 2297–2306. http://doi.acm.org/10.1145/2207676.2208388
15. David R. Flatla, Katharina Reinecke, Carl Gutwin, and Krzysztof Z. Gajos. 2013. SPRWEB: Preserving Subjective Responses to Website Colour Schemes Through Automatic Recolouring. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. 2805–2806. http://doi.acm.org/10.1145/2468356.2479521
16. Rui Gong, Haisong Xu, Binyu Wang, and Ming Ronnier Luo. 2012. Image quality evaluation for smart-phone displays at lighting levels of indoor and outdoor conditions. *Optical Engineering* 51, 8 (2012), 084001–1.
17. Legrand H. Hardy, Gertrude Rand, and M. Catherine Rittler. 1954. HRR Polychromatic Plates. *Journal of the Optical Society of America* 44, 7 (1954), 509–521.
18. Lane Harrison, Katharina Reinecke, and Remco Chang. 2015. Infographic Aesthetics: Designing for the First Impression. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1187–1190. http://doi.acm.org/10.1145/2702123.2702545
19. Ken Hinckley, Jeff Pierce, Mike Sinclair, and Eric Horvitz. 2000. Sensing Techniques for Mobile Interaction. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology (UIST '00)*. 91–100. http://doi.acm.org/10.1145/354401.354417
20. Shinobu Ishihara. 1917. Tests for colour-blindness. Pseudoisochromatic Plates. (1917).
21. Julie A. Jacko, Max A. Dixon, Robert H. Rosa, Jr., Ingrid U. Scott, and Charles J. Pappas. 1999. Visual Profiles: A Critical Component of Universal Access. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. 330–337. http://doi.acm.org/10.1145/302979.303105
22. Jaejeung Kim, Sergey Leksikov, Punyotai Thamjamrassri, Uichin Lee, and Hyeon-Jeong Suk. 2015. CrowdColor: Crowdsourcing Color Perceptions Using Mobile Devices. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. 478–483. http://doi.acm.org/10.1145/2785830.2785887
23. Youn Jin Kim, M. Ronnier Luo, Wonhee Choe, Hong Suk Kim, Seung Ok Park, Yeseul Baek, Peter Rhodes, Seongdeok Lee, and Chang Yeong Kim. 2008. Factors affecting the psychophysical image quality evaluation of mobile phone displays: the case of transmissive liquid-crystal displays. *Journal of the Optical Society of America* 25, 9 (Sep 2008), 2215–2222.

24. Kenneth Knoblauch, François Vital-Durand, and John L Barbur. 2001. Variation of chromatic sensitivity across the life span. *Vision research* 41, 1 (2001), 23–36.

25. Giovane R Kuhn, Manuel M Oliveira, and Leandro AF Fernandes. 2008. An efficient naturalness-preserving image-recoloring method for dichromats. *Visualization and Computer Graphics, IEEE Transactions on* 14, 6 (2008), 1747–1754.

26. Olof Lagerlöf. 1982. Tricyclic psychopharmaca and colour vision. *Documenta Ophthalmologica Proceedings Series* (1982).

27. Po-Hung Lin and Wen-Hung Kuo. 2011. Image Quality of a Mobile Display under Different Illuminations. *Perceptual and Motor Skills* 113, 1 (2011), 215–228.

28. Po-Hung Lin and Patrick Patterson. 2012. Investigation of perceived image quality and colourfulness in mobile displays for different cultures, ambient illumination, and resolution. *Ergonomics* 55, 12 (2012), 1502–1512.

29. Peter Liu, Fahad Zafar, and Aldo Badano. 2014. The effect of ambient illumination on handheld display image quality. *Journal of digital imaging* 27, 1 (2014), 12–18.

30. Ian J Murray, Neil RA Parry, Declan J McKeefry, and Athanasios Panorgias. 2012. Sex-related differences in peripheral human color vision: a color matching study. *Journal of vision* 12, 1 (2012), 18.

31. Allen B. Poirson and Brian A. Wandell. 1990. The Ellipsoidal Representation of Spectral Sensitivity. *Vision Research* 30, 4 (1990), 647–652.

32. Benedict C. Regan, J.P. Reffin, and John D. Mollon. 1994. Luminance Noise and the Rapid Determination of Discrimination Ellipses in Colour Deficiency. *Vision Research* 34, 10 (May 1994), 1279–1299.

33. Katharina Reinecke and Krzysztof Z. Gajos. 2014. Quantifying Visual Preferences Around the World. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. 11–20. http://doi.acm.org/10.1145/2556288.2557052

34. Maureen C. Stone. 2003. *A Field Guide to Digital Color*. A. K. Peters, Natick, Massachusetts, USA.

35. Françoise Viénot, Hans Brettel, and John D. Mollon. 1999. Digital Video Colourmaps for Checking the Legibility of Displays by Dichromats. *Color: Research and Applications* 24, 4 (1999), 243–252.

36. Colin Ware. 2012. *Information visualization: perception for design*. Elsevier.

37. J. Terry Yates, Ioannis Diamantopoulos, and Franz-Josef Daumann. 2001. Acquired (transient and permanent) colour vision disorders. *Operational Colour Vision in the Modern Aviation Environment, NATO RTO Technical Report* 16 (2001), 43–47.