https://doi.org/10.2352/ISSN.2470-1173.2021.16.COLOR-308

Examining Spatial Attributes for Color Deficient Observers*

R. Eschbach[1,2] and P. Nussbaum[1]

[1] Norwegian Colour and Visual Computing Laboratory, NTNU - Norwegian University of Science and Technology , Gjøvik, Norway
[2] Monroe Community College, Rochester, NY

Abstract:

Color deficiency is a common affliction for males, with about 8% of the male population suffering from some form of – colloquial – color blindness. In common description, this is interpreted as the inability to perceive certain colors from a sensory standpoint. If that description is sufficiently complete, spatial attributes of the color signal should not influence the visual experience of a Color Deficient Observer (CDO).

In this paper we describe the set-up and results of an experiment that is intended to vary the spatial attributes of a color deficiency test when shown to an observer. For this, we modified color vision charts to maintain their colorimetric attributes while varying spatial attributes and we measured color deficient observers with respect to their response accuracy and their response speed. We show results that indicate that there is a clear influence of color proximity in the responses.

Introduction:

Color Deficiency is a rather common occurrence in an average population, with about 8% of males being effected. Despite the prevalence of the problem, it is often overlooked, ignored or simply misunderstood. This is obvious in the development of graphic designs in orientation guides (e.g.: color codes inside hospitals), in education (Ref [1]), in instrument design (e.g.: indication lights), computer interfaces and security situations. Here, the awareness to address the issues and problems faced by color deficient observers is generally absent.

A good example in the education field can be found in Ref[1] where - in an early Math Schoolbook - the student is asked to count the circles of a specific color to define fractions. This is reproduced (in part) in Figure 1.

In this situation, a color-deficient student is easily mis-classified as math-deficient with the typical negative consequences.

Another example is a photo taken inside a train (Fig 2) indicating the method to escape in case of an emergency. To a color deficient person, the red-dot might be invisible, despite being an important safety item.

* This work was supported by the Regional Research Funds (RFF) Inland Norway

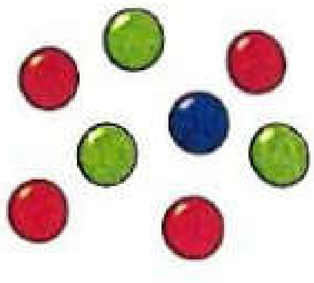

Figure 1 Figure 1: A problem in a math book teaching fractions and asking the student to count the number of red, green, blue circles. Ref.[1]

Figure 2. Emergency exit sign in public transportation, taken by one of the authors.

This paper describes an experiment performed to better understand the connection between color and the spatial structure of the color. It is a continuation and refinement of an early exploratory paper given in Ref[2]. There it was postulated that the performance of color deficient observers on common color vision tasks varies with the spatial structure and amount of edge information. In this paper we will show the experimental set-up and share the data obtained (strongly restricted by the COVID situation).

Color Deficiency Test:

Color deficiency tests are designed to probe the color discrimination of the test-observer. A good overview of tests can be found at Colblindor [3]. Color deficiency tests normally consist of separated colors, not containing any edges or not varying the spatial relationship of colors. This is, of course, a proper approach to assure that pure color discrimination is tested and not some other effect or artifact. In our case, we intentionally want to introduce color edges since we are trying to examine the interplay between pure color perception and the perception of colors in the presence of edges or neighboring colors. At the same time, we intended to stay close to a standard test in order to avoid the complexity of natural scenes.

We decided to follow the general structure of Ishihara charts since in their normal form, all colors are separated by a defined background color and by changing the size of the color circles one is also able to change the edge content. Our Ishihara-like charts were self-created and were not identical to the actual Ishihara test. Ishihara is part of a group called “plate tests”, since the observer is asked to identify an object (often a number or simple symbol) that differs from the background only by its color.

One of our charts is shown in Figure 3. Since the charts are described as vector data it is easy to modify the sizes and other attributes of the circles in a predetermined way.

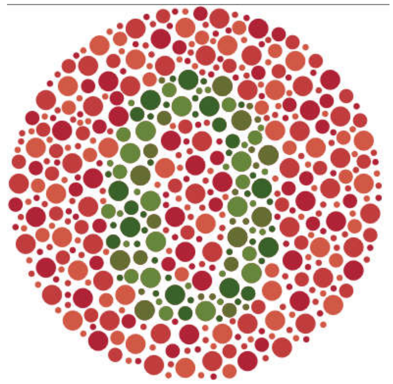

Figure 3. Color chart used as one of the basic charts in the experiment.

Color Confusion Lines:

One basis for color deficiency tests is the existence of confusion lines, i.e.: lines in the xy diagram that cross different colors to a color normal observer, but lead to the “same” color perception of a color deficient observer [4,5]. The confusion lines connect a single point (copunct) to a spectral color and colors along those lines should be indistinguishable to a color blind observer. In our case, it is important to note that we are interested in color deficient observers, meaning observer that do not lack a cone-type, but that have other anomalies, like response shift, spatial distribution, etc. For these people, the likelihood of “seeing” a color is noticeably reduced. And, as a general assumption, they should make more mistakes in identifying a color and take more time to perform a color task.

A schematic drawing of the CIE diagram with a set of confusion lines and the copunct (CO) is given in Figure 4.

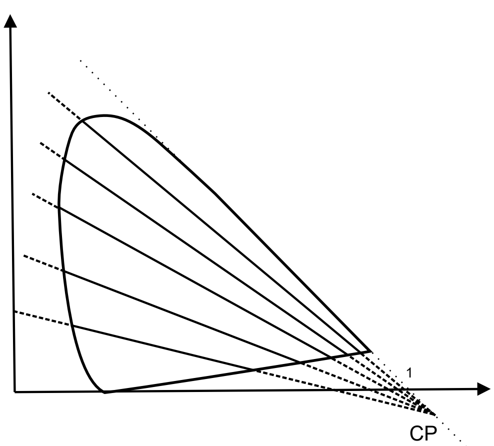

Figure 4. A conceptual drawing of confusion lines (solid) emanating from the copunct (CP) and traversing the visible spectrum.

Literature has a set of copunct location varying around (0.74, 0.26) for protanopes and varying more expressly for deuteranopes from (1, 0) to almost (2,-1) (see e.g.: p 1320 in reference[6])

Experimental Design:

The hypothesis that we are trying to test in our experiment can be formulated as follows:

> *Does the proximity of colors of in a plate test influence the speed and/or accuracy of the object detection for a color deficient observer ?*

Emphasis here is on the color deficient observer, meaning the observer is identified by a standard plate test (like Ishihara) but does not necessarily completely lack one of the cone types. We are looking for the performance of protanomalous/deuteranomalous observers, not only protanope/deuteranope observers. Our hypothesis should be testable by measuring accuracy of recognition and speed of recognition for at least some of the color charts.

Since we are trying to see a change in the performance and not the absolute performance of the observer, we picked one copunct location within the common locations for a deuteranope observer (1.43, -0.43).

In order to cover the entire range of colors, we picked one confusion line to run through gray (0.33, 0.33) and the remaining 9 lines to connect to the spectral points at equal intervals from 495 to 515nm[2], giving us a total of 10 different confusion lines.

With ten different symbols (numerals from 0 to 9), we had 100 potential plates. Of these, we actually created 31 files (each symbol appeared at least 3 times, each confusion line was used at least 3 times) as starting point for the experimental data. All data was contained within the sRGB gamut to avoid further processing. These plates had the "standard" size and all other charts were derived from these charts. All the plates were created as scalable vector graphic allowing us a simple and efficient way to change the color description of the individual circles, as well as the sizes of the circles.

Spatial Attributes:

The basic assumption of this paper is that a change in spatial attributes of the charts without changing the physical color components will influence the ability of a color deficient observer to correctly identify the symbols inside the charts.

For this type of charts, it is easy to modify the spatial relationship / edges of the color dots without changing their actual color values. In order to change the proximity or edge content of the charts, we varied the dot size of the individual dots inside the chart. This is shown in Figure 5.

The standard plate (left most segment in Figure 5.) contained circles of three different sizes (s, m, l) and the increase in circle size was achieved by scaling the circles, In order to maintain consistency for the different plates, we decided to change the radius of the different circle by 5% (large circles), 13% (middle circles) and 30% (small circles). This scaling was chosen so that the final chart was not completely dominated by the original large circles. Overall, three scalings were performed, giving us 124 charts that could be used in the experiment.

[2] Some wavelengths were rounded to the next integer nanometer number.

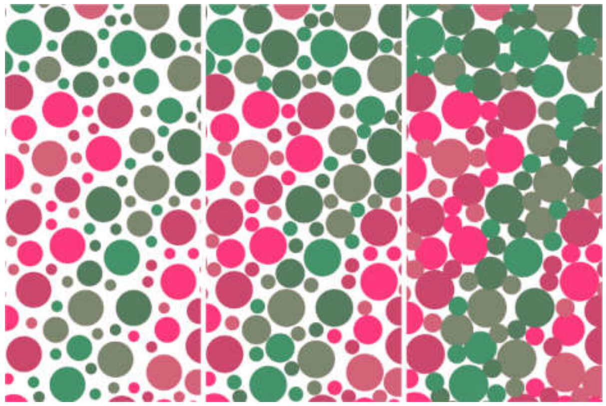

Figure 5. Segment of one color test plate showing the variation in circle size to create proximity and edges.

Experimental Data Creation:

In the originally intended experimental design, we expected to identify color deficient observers and to invite them into the lab to perform the experiment under controlled lab conditions. The COVID situation, however, forced a redesign of the experimental data gathering. Identifying observers and bringing them into the lab was deemed impossible. We thus decided to create a “stand alone“ software that could run without supervision. This also meant, that we could not – or not easily – control who took the test and thus would likely get a large number of color normal observers. We thus changed the experimental layout and created a program following the flowchart of Figure 6.

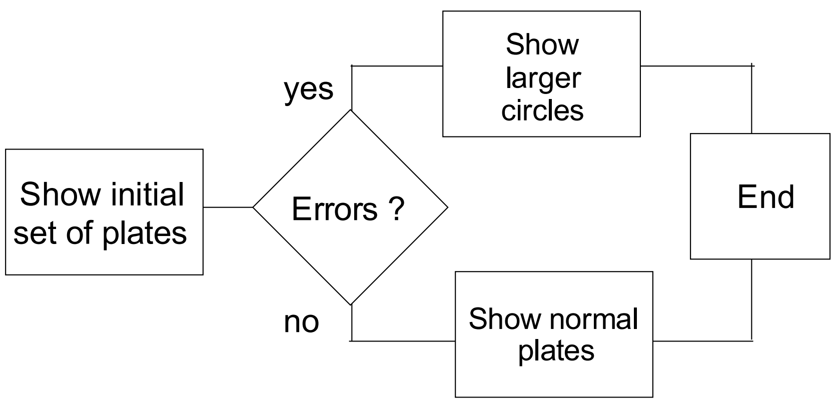

Figure 6. Flowchart of the standalone software identifying color deficient observers

The standalone software had to address several issues that would not occur in a lab setting:

(1) It had to be self-explanatory.
(2) It had to be attracting observers without human input, this includes the look of the welcoming page, as well as the opinion given by observers afterwards
(3) It had to be careful with the collection of personal data

In our base case, the software was installed on a laptop and that laptop was located in a corner of a large lab room where a limited number of students was present at any point in time due to COVID restrictions. The set-up is thus an example of an “uncontrolled” visual experiment where additional noise factors could enter into the measured data. However, this experimental noise should not influence the general trends, since we are looking at the relative performance of the observers in the (for the individual) constant setting. At the start of the experiment, a random set of standard plates was shown (picked form the original circle sizes) and the program calculated the error rate of the observer. If at least 2 of the first 7 plates was misclassified (either a wrong answer entered or no symbol seen), the observer would follow the path that shows increased circle sizes. Again, the different sizes were mixed, interspersing standard circle sizes with increased sizes. A total of 25 plates were tested for each observer. Observers that had passed the initial test were led down a different path, where smaller circles were shown to maintain interest. The intention was to make it look progressively more difficult, so that normal observers would not discourage others from taking the test.

For observers that were led down the color deficient path, the 25 plates consisted of roughly equal numbers for the different sizes. Since we had 4 different circle sizes (1 standard and 3 enlarged), the observers saw approximately 6 of every chart type. Observers taken down this path also received the request to leave contact information completely voluntarily in the form of an e-mail address.

In some instances, observers were asked to participate directly and although one of the authors was present, the same standalone software was used for consistency. This happened for people with known or suspected color deficiency.

It should be noted that the availability of a completely standalone software allowed us to approach locations for a quasi-permanent installation of the software, with one set-up being installed at time of writing in a Science Museum[3] in Norway and one in the University Library of NTNU Gjøvik.

Collected Data[4] (ongoing):

Despite the problem with data collection during times of limited access to the universities, a total of 10 observers were identified as color deficient and some of them were actually tested a second time with a supervised test and also asked to perform the Munsell 100 Hue test. Again, this was strongly limited by the current situation. One of the authors (Eschbach) is severely color deficient. He did not participate in the tests and no data of his is included in the analysis of this paper.

Setting the threshold for being classified as color deficient to missing 2 out of the first 7 tests was an arbitrary number, but we did not want to classify people that made a single mistake, maybe by hitting a „9“ instead of „0“ when the keys were close together. Also, there might be a learning effect with an initial mistake being made as part of understanding the test. A seconday filtering could have been performed on the data, but in the data presented here, all observers that took the „color deficient path“ are included. This likely makes the data actually look worse in terms of hypothesis testing.

The program used to create the data was written in Python starting from a PsychoPy (Ref. [7]) framework that allows the collection of precise timing data.

[3] We are expecting a large number of minors in the Science Center installation and thus do not even ask for voluntary information like e-mail or name. Only the raw data is collected.

[4] One of the authors (Eschbach) is severely color deficient. He did not participate in the tests and no data of his is included in the analysis given in this paper.

The combined result for all 10 observers is given in Table 1, where Chart Type „1“ refers to the standard size (left in Figure 5), whereas „2“ (middle in Fig. 5), „3“ (right in Fig. 5) and „4“ (not shown) refer to the three chart sets with the enlarged circles.

Chart Type	Correct Observations	Total Observations	<Time for correct>
1 (standard)	71	163	3.33[s]
2 (larger)	55	95	3.41[s]
3 (larger)	72	95	2.71[s]
4 (larger)	108	132	2.52[s]

Table 1. Collective result for all 10 observers considered color deficient.

The difference in the charts tested for each circle size is caused by the random selection of the large pool of charts. Since we used the standard circle size at the beginning to identify color deficient observers, the number for that size should be (and is) larger than the other numbers.

From Table 1, the hypothesis that we started out (originally proposed in Ref.[2]) seems to be confirmed. We can also look at the results in pure terms of percentages as is done in Figure 7(a) and time for correct answers (b).

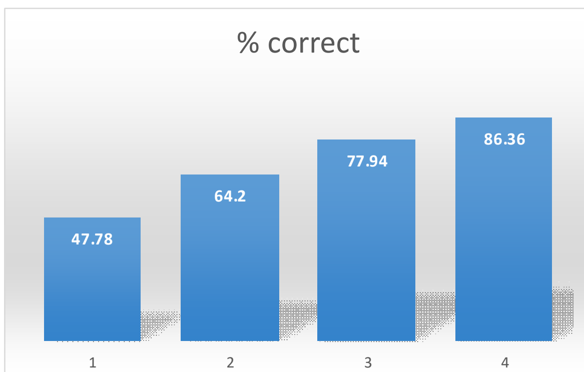

Figure 7a. Percentage of plates that were correctly identified.

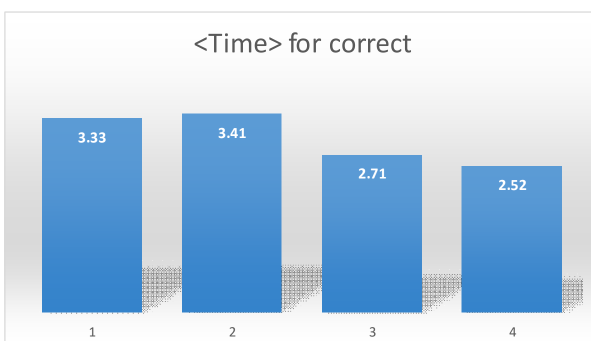

Figure 7b. Average time for the observers to correctly identify a symbol. Times for skipping the plate or answering wrongly were not used.

The data for correctly identifying the symbol encoded inside the plate increases from 48% to 86%, essentially doubling. This is in alignment with the preliminary data from Ref.[2], however, there only a small number of known color deficient observers were tested.

The data for the time it took to identify the symbol correctly needs to be taken with care, since there are several effects that likely will introduce unwanted variations for a small number of observers (10 in our case). For example, there seems to be a clear case of “learning”, thus the times for the standard charts, seen first, might likely be larger than actual numbers (for this small set of observers we did not filter chart order which is available from within the program).

Also, basic questions about the timing measurement exist. Should only the correct answers be considered ? What should be done to differentiate a quick “skip” from a serious attempt because the observer “almost” can identify the symbol. Our current assumption is that with an increased number of measurements, this problem (hopefully) alleviates itself.

Follow-up with observers:

There are various reasons to attempt to follow-up with a color deficient observer. The most obvious one is the ability to create a larger data set by testing them more thoroughly. A second reason is to create comparative evaluations using different color tests. In our case, we have chosen to ask the observers about taking the Munsell 100 Hue test. It should be noted, that all observers identified as color deficient chose to voluntarily give their contact data, but only a few underwent further testing in the lab.

Figure 8 shows the Munsell 100 Hue results for two of the observers that were available for a follow-up. Observer #1 had a correct rate for the standard charts of 77%. Observer #6 of 25%. For the largest circle size, Observer #1 received a 100% correct rate, Observer #6 of 77%. The Munsell data correlate well with our findings.

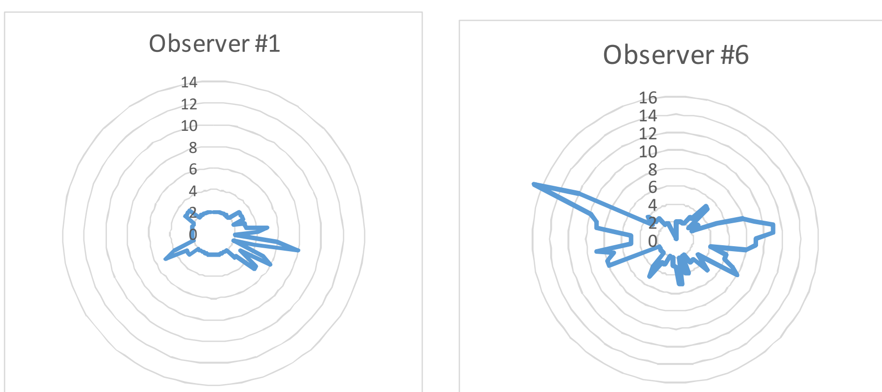

Figure 8. Munsell 100 test results for two observers.

Summary:

We have conducted a limited number of color deficiency tests attempting to see a relationship between the recognition rate of an Ishihara-type color test and the proximity / edges contained in the test. The hypothesis that there is an interaction, enabling color deficient people to start seeing the symbols contained inside the pattern of the color plates is supported by the data generated from 10 color deficient observers.

The experiment was done predominantly in a standalone fashion and we tried to contact the color deficient observers after the analysis of the test results. This was largely successful. Of the people taking the test in the standalone version – simply because the test was in the room and others had taken it – one student later declared that he had not known about his color deficiency before. That student had missed 43% of the standard charts, missed 14% of the simple size increase, but correctly answered all larger circle plates. Due to the current situation, no follow-up could be performed.

Conclusion:

Color deficiency tests are useful tools in identifying color deficient observers. Often these tests are used to make statements about "how color deficient observers see". This extrapolation, however, is more complicated that simple models imply. Our experiments attempt to create the supporting data to define the relationship between color and spatial attributes for the ability to distinguish colors. This information hopefully will influence the future of vision models that simulate color deficient observers.

The current data set seems to support the hypothesis that edge content is an important feature. However, the number of color deficient observers in the tests was small with N=10 and no in-depth testing could be performed on the majority of the observers.

References:

[1] A. K. Kvitle, "Should colour vision deficiency be a recognized special education need", *doi:10.3233/978-1-61499-923-2-832*

[2] R. Eschbach, S. C. Morgana, A. Quaranta, C. Bonanomi, A. Rizzi, "Feeling edgy about color blindness", Dark Side of Color, Color Imaging XIX, Electronic Imaging 2014.

[3] Color Blindness Tests, Colorblindor, https://www.color-blindness.com/color-blindness-tests/.

[4] G. A. Fry, "Dichromatic color confusion lines and color vision models", J. Optometry & Physiological Optics, pp933-940, 63(12), 1986.

[5] D. B. Judd, "Standard response functions for protanopic and deuteranopic vision", Journal Nat. Bureau of Standards, pp407-437,(33), 1944.

[6] D. B. Judd, "Fundamental studies of color vision from 1860 to 1960", Proc. N. A. S. 1313-1330,55(6), 1966.

[7] PsychoPy is available at www.psychopy.org