An Analysis of Color Discrimination of University Students in the Graphic Arts

Brian P. Lawler

Keywords: color, vision, physiology, students

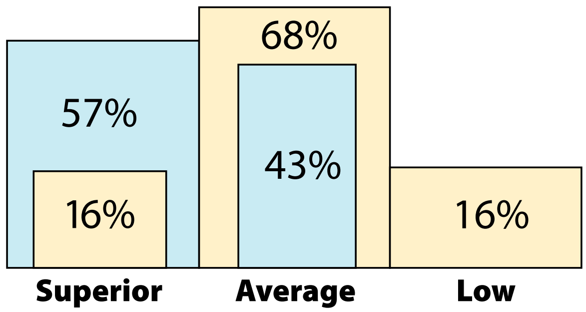
Abstract:
Since 1999 I have been teaching a course in Color Management and Quality Assessment as part of the baccalaureate curriculum in Graphic Communication at Cal Poly, San Luis Obispo. In the years that I have taught the course, I have had nearly 900 students, and I have required every one of them to take the Farnsworth-Munsell color discrimination test as well as the Ishihara color-blindness test.

In the early years, I scored the tests by hand, using a scoring sheet provided by the manufacturer. Though that was successful, I was never able to correlate the data from a number of tests.

In 2007 and 2008 I entered the students' scores into the electronic scoring version of the Farnsworth-Munsell color discrimination test, and was able for the first time to make sound observations about the scores of the student group. At the end of the Fall quarter, 2008, I went back to the paper records of my former students and began entering the data from those courses. This paper includes students from Winter, 2005 to Fall of 2008, a total of 218 students. 156 of those students are or were enrolled in Graphic Communication, while the balance are my control group. The control group consists of students who are specifically *not* enrolled in the Graphic Communication program. The control group students are in many majors including Animal Science, Industrial Technology and Engineering. Though I did not encounter any, I had planned to eliminate from the control group any students in Art, Graphic Design or Architecture.

Brian P. Lawler is Assistant Professor of Graphic Communication
at California Polytechnic State University, San Luis Obispo, California.

The results of my analysis of scores on the Farnsworth-Munsell test were intriguing to me, even in the early years when I scored the tests manually. Graphic Communication students exceed the norm. According to the Farnsworth-Munsell test instructions and accompanying documentation, 16 percent of the general population score with *Superior* color discrimination. More than half of my students score in this range. And, right there is the basis for this research. *Why do students in our graphic arts program perform better in color discrimination tests than average people?*



A large percentage – 57% – of students of Graphic Communication have *Superior* color discrimination. The general population* has just 16% who qualify as *Superior*. Those students with *Average* color discrimination fall a little below the general population, none scored with *Low* color discrimination.

The Farnsworth-Munsell 100-hue Test

The test is a set of four boxes of plastic cylinders with color caps. It tests a person's ability to *discriminate* between colors – to put them into order. Though entitled *100-hue,* the test actually has only 85 cylinders. According to Munsell scientists, the color discrimination test has been changed several times to simplify it.

Before taking the test, the cylinders are rearranged randomly by a test ad-

---

*Studies of the general population are published by Munsell Color, a division of X-Rite. General population data are taken from publications that accompany the Farnsworth-Munsell 100-hue Test. Farnsworth-Munsell discounts tested color-blind subjects from their analyses; as a result, I also discounted proven color-blind subjects from my analysis. I identified one student with such color-defective vision in this group of participants.

ministrator (they are numbered on the bottom), then the test is given to the participant.

The person taking the test is then asked to rearrange the color-topped cylinders into color order between the two fixed end cylinders. The test instructions suggest – but do not require – giving the person two minutes to take each of the four parts of the test, but also suggest that time is not a factor in success. Since I did not originally require timeliness on the test, I have continued in the same vein for all of my students. Most participants typically complete the entire test within 15 minutes.

When the test is complete, the administrator closes the boxes, turns them over, then reopens the boxes upside-down, recording the results as a series of numbers on the scoring sheet. Each transposition of colors is recorded as a line over a pair of numbers. More severe transpositions are noted as a series of numbers written above the score sheet numerals. That information is entered into the Farnsworth-Munsell (hereinafter *F-M*) test scoring software by dragging the on-screen cylinders into the same positions as the test results. Student name, the date of the test, and the gender of the participant are also entered, as are experience level, location, and whether the test results are first-time or subsequent takings of the test.

I recently added *date of birth* to the information I am collecting, as age *may* be a factor in these results (F-M makes no assertion about age other than to say that color discrimination may get worse as we get older). But, it will be several years before I can make any reasonable analysis based on the age of my test-takers. I will always be testing students in the 21-25 year-old range – the age of college students.

As data are entered into the F-M scoring software, a database is created. Each entry is one line of the database, with some analysis possible, including the individual's score or combined – or averaged – scores of a group of participants.

The F-M scoring software also allows the data to be exported to a text file that is compatible with Microsoft Excel and other analysis software.

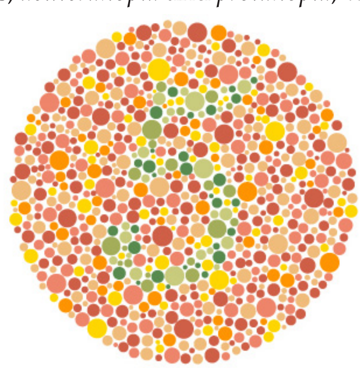
Lighting conditions recommended are for daylight-balanced light of D65. My testing environment is D50, but I chose to use it over the years as it is measurable, consistent and seldom-used in our laboratories, which means that the light source is a constant. The booth is a GTI model with an indicator of light quality and warm-up period, both of which I observe carefully. The color temperature difference is significant, and may be a factor in the common transposition of cylinders 28 and 29, discussed later.

Testing conditions

I have been strict in testing my student subjects, requiring each of them to use the same test kit, and to view the color cylinders in the same lighting booth. I have never changed the test environment, nor have I moved or changed the circumstances of the test area. Lighting and test materials are, as a result, a fixed value in this large study. Considering that I have been conducting the test for ten years, it represents a significant body of color discrimination data, perhaps the largest in modern times.

I am unaware of any similar collection of data (and F-M scientists have confirmed this). Other similar tests have been carried out with some variability – changes in test location, lighting and other environmental conditions that do not affect my test results.

In addition to testing for color discrimination using the F-M 100-hue test, I also have each student take the more common *Ishihara Color Blindness Test.* This test, originally created by Dr. Shinobu Ishihara (1879-1963) in 1918, tests for a person's ability to see color. It tests for two of the most common color defects, *deuteranopia* and *protanopia*, which are red-green vi-



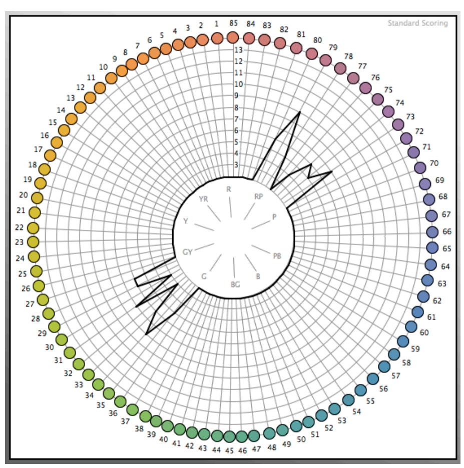
This plate from the *Ishihara Color Blindness Test* is designed to identify red-green color-defective vision. If you don't see the number 5 in the plate above, you may have a color defect.

sion defects. Dr. Ishihara's test does not test for the much rarer form of color blindness called *tritanopia,* which is a blue defect resulting in near-total color blindness. I suspect that Dr. Ishihara may not have encountered people with that form of color blindness, but there is no documentation on this topic that I have been able to find.

These three forms of color-defective vision are strictly genetic, though some people lose their color sight as a result of injury or disease.

My reason for testing with Ishihara is to show the students that there is a difference between color *blindness* and color *discrimination.* It is *possible* (but unlikely) that a person who has color-defective vision may still be able to discriminate the colors in the F-M test by virtue of their grayness or tone. That possibility is addressed by the F-M documentation, and also by my experience. In the ten years of giving these tests, I have encountered two profoundly color-defective students and a number of students who show some red-green color defects. The two color-defective students are brothers, and one did not know that he was color-blind. The younger brother insisted that his brother was not color-blind. Their father is not color-blind (color-defective genes are dominant Y), and he had no knowledge that either son has color-defective vision.

Students who fail to identify some of the targets in the Ishihara Color Blindness Test usually display that same color defect in taking the F-M test. The test will show two strong spokes in opposite directions on the scoring sheet as shown in the illustration below (simulated).



I have also encountered anomalous test results on occasions when students were self-declared as sick or very tired. When this occurs, I discard the results rather than to re-test the student. Re-testing creates another variable that I have chosen not to pursue, as Farnsworth-Munsell indicate (and my experience proves) that people who take the test a second time perform significantly better. To allow a student to re-take the test would require me to retest *all* the students, and time does not allow this.

At the conclusion of the testing period, I produce an individual color vision report for each student, a two-page explanation of their test score with some generic information about the two tests and a screen-capture of their individual test score. I keep this private so as not to embarrass anyone who is sensitive about their ability to see or not see color.

If they have responded to Ishihara with errors, or show a significant error on F-M, I discuss it with them personally, and ask if they are aware of their color vision. Most respond that they are keenly aware that they have trouble with red-green color identification, though some are surprised by the test results. It is interesting to me that many of the students have never taken the Ishihara test in their lifetimes. I first saw this test when I was in elementary school, and had always assumed that it was a normal part of testing children in American schools. That is obviously not true.

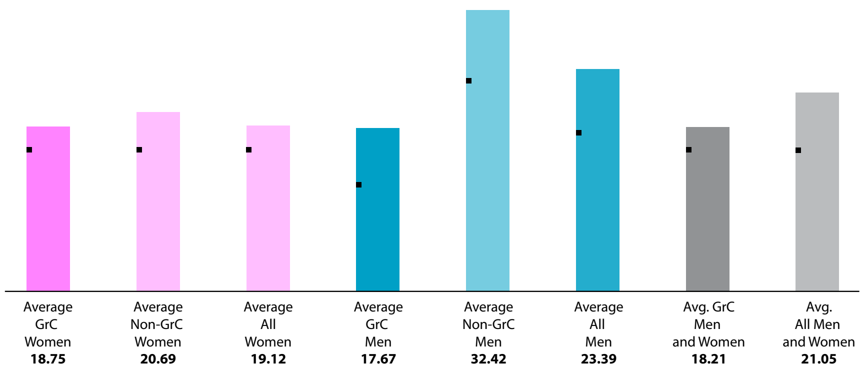
Test scores

A perfect score on F-M is zero, and this is extremely rare. In my population of student participants, a total of eight scored perfectly, six male and two female. All of the participants who scored perfectly are Graphic Communication (hereinafter "GrC") students.

Scores of *Superior* color discrimination were made by 66.6 percent of GrC men, and 51 percent of GrC women. In the control group, the men scored in this Superior category 36.8 percent, while the female control group outscored their GrC counterparts by a slight margin – 52 percent. [My female control group is not adequate for a statistically significant analysis; I have only 23 control group females. In the coming years I will work to get more students outside my department to participate in the control group. This is quite difficult.]

Scores of *Average* color discrimination were made by 33 percent of GrC men and 60 percent of the control group men. These, of course, are the inverse of their performance in the *Superior* category as there are no students in the GrC category who scored in the *Low* category of discrimination, and only one male control student who scored in the *Low* category.

Of the females, 48.96 percent of GrC women score in the *Average* category,



Scores for the participants in my study are broken into male and female, GrC and non-GrC students. The small black squares indicate the Median for each category. A perfect score is zero.

and 47.83 percent of the control group women score in this category. None of the women scored in the *Low* category.

If I contrast these numbers to those published by Munsell Color, the glaring contrast arises. While the students' scores are generally excellent, they are amazingly high when compared to the population at-large.

Munsell Color says that 16 percent of the general population scores in the *Superior* category, and that 16 percent of the population scores in the *Low* category, the balance being in the middle – *Average* – range. But, 57 percent of tested GrC students score in the *Superior* category, and the rest score as *Average*. None score as *Low*.

With very few exceptions, the students who have taken this test under my supervision have not seen the test before, and those who have are usually only vaguely aware of it. Some have taken the online version of the test which X-Rite had on their web site (it was available for a while in 2008, quite well hidden, and has now disappeared from the site with no mention of its existence). The online version does not quantify the score in the same way as the physical test does, and is subject to the vagaries of the individual's computer display, which is probably why X-Rite has taken it down. Without certification of the quality of color on the receiving end of the process, there is no guarantee of anything close to accurate color.

Why?

For ten years I have been fascinated by these statistics. At first I thought they were anomalous, and would slide eventually to the norm, but as time went on, the statistics became more pronounced, not less, and that observation led me to write this paper.

If you flip a coin enough times, the long-term average of heads and tails becomes 50:50. I figured that if I tested enough students, eventually the averages would approach those of the general public. That never happened.

Graphic Communication attracts people who are interested in print, photography, page layout, design, typography, *and color.* These students are typically creative, but they are also not students of pure art. In fact, at Cal Poly, Graphic Communication is considered a discipline for those interested in a career in management of printing companies and the services that feed the printing industries.

Though we have a *Design Reproduction Technology* concentration in our program, it is equal parts art and production. It includes prepress, troubleshooting, knowledge of file formats, color management and imposition. It is not pure art, and does not usually attract the *pure artist.*

Left-brain/Right-brain?

Our students do equal parts theory and hands-on production. They are involved in the actual printing, folding and binding of printed products, but they also study costing, estimating, marketing, color theory, quality assessment and human resources management. Clearly, this is not an art school.

A mathematician friend suggested to me that he is as good with numbers as my students are good with color. He suggested the right-brain idea, and though I think it is possibly a factor, it is certainly not the *causative* factor. Creative people are not necessarily capable of discriminating colors well.

So, why are students attracted to the GrC major? More specifically, why would students who have *Superior color discrimination* choose to study the graphic arts? A corollary question might be asked about students in mathematics and science: is the color discrimination of that population of students *Average* or *Low?* And, is there any correlation between one's chosen college program and one's ability to discriminate between colors?

My control group is comprised of students from many disciplines, so the results of their tests should be telling in this respect. They outperform the general population also, though their scores reveal that they are (with the exception of the female control group) not as good at color discrimination than their GrC counterparts. More could be learned by testing a larger number of non-GrC students, which I intend to do in the coming years.

Is color discrimination *learned?*
Farnsworth-Munsell says yes – to a point. If a person takes the F-M test, and then re-takes the test, their score will typically improve by a factor of 30 percent, but additional passes result in no significant score improvement. Realistically, one can only do so well at the color discrimination test, a result of genetically-defined physiology and a developed skill for telling hues apart.

Certainly there is a factor in color discrimination for *working with color.* As with any skill, if you do it long enough, you get better at it. With color, there is probably a level of skill that can be attained by continuous application of color and the making of color decisions. The counter to this, though, is that genetic factors can hamper some people, making it difficult or impossible to "learn" how to discriminate between different hues.

How much of color discrimination is age-related?
Does a person's color discrimination ability change as they get older? The information Munsell provides on their general population studies indicates that their samples come from a wide range of ages – 15 to 45. It is unlikely that our color vision is so much better in our 20s as to skew test results this much. Some other factor must be at work here.

Barring disease or injury, our color perception stays constant for our entire lives. We might not be able to focus, but we can certainly put colors in order.

Color discrimination as a job function

There are many people who make their living selecting, analyzing and adjusting colors. These people decide what colors will be used to paint automobiles and park benches. There are people who compare colors to master patches and give a thumbs-up or down to products as they come off an assembly line. These jobs require the worker to be able to discriminate colors correctly and consistently.

Screening candidates for jobs that require excellent color vision, and excellent color discrimination requires the use of tests like the Ishihara and the Farnsworth-Munsell tests.

Curiously, color-defective vision is – so far – not considered a *disability.* People are born with their color vision, or lack thereof, and unlike other defined disabilities, testing poorly for color discrimination can be used by an employer as a reason to exclude a person from a job where such skill is required.

The mysterious cylinders 28-29

In many cases those who transposed only one pair of color cylinders did so by transposing numbers 28 and 29.

At first I thought this was due to smudges on those cylinders, dirt or fading. After re-reading the F-M instructions, I have come to the conclusion that the use of a D50 light booth is the most likely source of the problem. My tests have been conducted in an environment that is incorrect for the test as it is designed.

F-M sells replacement cylinders for a few dollars each, but I fear that changing the cylinders after testing students for ten years would introduce a new variable into the process that would potentially invalidate the test scores before or after the change. I have chosen to leave the cylinders unchanged, and the light source unchanged.

Year 11 and beyond

As I continue my testing this year I will evaluate the color discrimination of the students who pass through my classes as I have in the past. And, I will continue to search for an answer to the question of why these students score significantly better than the general population.

I will also continue to enter the test scores for the hundreds of other students who have taken the test, but whose scores are not yet in my database.

Perhaps if I flip the coin enough times the average will level-out to 50:50.

Brian P. Lawler
San Luis Obispo, California
February, 2009