Comments on the University of Joensuu's Matte Munsell Measurements

Paul Centore

Abstract

The University of Joensuu's measurements of the 1976 Munsell Book are one of the few publicly available sets of spectral reflectance measurements of Munsell samples. This article evaluates colour differences between the printed Munsell matte samples and the 1943 renotation standard. The median difference was found to be 3.6 or 6.9 CIEDE2000 units, depending on which spectrophotometer was used for the measurements. Furthermore, the Munsell values of the printed samples showed strong biases relative to the renotation, as did some of the hues and chromas. By comparison, the median differences for two Munsell books produced after 2000 are between 1.5 and 2.0, and all the individual Munsell components are bias-free. Their deviations from the standard suggest that Joensuu's measurements are not a sound basis for further analysis of the Munsell system.

1 Introduction

The Munsell system is a common and long-established colour specification system. Albert Munsell devised it in the early 1900s, as a perceptual system. In 1943, the Munsell renotation[1] reformulated the Munsell system quantitatively. The renotation was based on a comprehensive set of empirical judgements by human observers, of a wide gamut of painted colour samples, and is today the standard for the Munsell system.

In addition to a rigorous definition of the Munsell system, practitioners need physical exemplifications, such as books and posters, containing printed samples or swatches. Over the years, several such books have been produced.[2,3,4] Pre-eminent among these are the books produced by the Munsell Corporation itself. Through a series of purchases, the Munsell Corporation is now owned by the corporation X-Rite, whose regularly produced Munsell books have become an industry standard.

To be useful, physical Munsell exemplifications should be accurate: i.e. they should agree with the renotation standard. To test accuracy, a sample's reflectance spectrum can be measured with a spectrophotometer. The sample's CIE coordinates can be calculated, under the assumption that Illuminant C is the light source. The renotation provides aimpoints for the sample's CIE coordinates, relative to Illuminant C. A colour difference equation, such as ΔE_{00},[5] can express the difference between the actual coordinates and the renotation aimpoint as a single number. The lower this number is, the more accurate a match is.

In 1988, Joensuu University of Finland measured the 1976 edition[2] of X-Rite's *Munsell Book of Color— Glossy Finish.* Two sets of measurements were made,[6] the first with a Perkin-Elmer Lambda 9 spectrophotometer, and the second with an Acousto-Optical Tunable Filter (AOTF). Both sets of measurements, which are freely available online,[7] produced reflectance spectra, which were mainly used in principal components analysis. As far as is known, no other measured reflectance spectra of Munsell colours are publicly available.

The current article analyzes the Joensuu Munsell measurements, in terms of the Munsell renotation. Differences were calculated, using the CIE ΔE_{00} expressions, between the printed colours and the renotation standard. In general, the accuracy is very poor: the median ΔE_{00} for the Perkin-Elmer spectrophotometer is about 3.6; for the AOTF it is about 6.9 . For comparison, a 2007 *Munsell Book of Color* and *The New Student Munsell Color Set, 2nd ed.*, were measured, with an X-Rite ColorMunki Design spectrophotometer. The median ΔE_{00}s of these two more recent publications were between 1.5 and 2.0, much better than the 1976 edition. It is unclear why the 1976 errors are so large. Possibly the printed samples changed colour between 1976 and 1988, or possibly the spectrophotometers in 1988 were more variable than today's spectrophotometers. In any event, the discrepancies and ΔE_{00}s suggest that conclusions based on the Joensuu measurements should be regarded as questionable.

2 Analysis Method

2.1 The Munsell Colour System

The Munsell colour system is a colour system that specifies surface colours by three perceptual attributes: hue (H), value (V) and chroma (C). The Munsell notation for a colour takes the form H V/C, or, for neutral (N) greys, it takes the form NV.

Hue is universally understood. It says whether a colour is red, yellow, purple, etc. Munsell designates 10 basic hues: R (red), YR (yellow-red, or orange), Y (yellow), GY (green-yellow), G (green), BG (blue-green), B (blue), PB (purple-blue), P (purple), and RP (red-purple). Each basic hue is further subdivided into 4 steps, denoted with a prefix. For example, the four greens are denoted 2.5G, 5G, 7.5G, and 10G. A prefix of 10 is sometimes replaced with a prefix of 0 and the next hue. For example, 10G is sometimes written 0BG. In all, the Munsell system specifies 40 hues which are equally spaced perceptually. White, black, and greys are not considered hues in the Munsell system. Many different colours can have the same hue. A set of colours of various values and chromas, but all with the same hue, is called a "hue leaf."

Munsell value designates how light or dark a colour is. The theoretically darkest black has a value of 0, and is denoted N0. The theoretically lightest white has a value of 10, and is denoted N10. Between N0 and N10 are 9 progressively lighter greys, denoted N1, N2, and so on up to N9. The spacing between the greys is perceptually equal.

Munsell chroma refers to how intense, or saturated, a colour is. A dull colour is closer to a neutral grey than an intense colour. The Munsell system denotes chroma numerically. Greys have chroma 0. A colour with a chroma of 10 or higher is generally perceived as saturated. Colours of low chroma, say 4 or less, are perceived as subdued, with a high grey content.

2.2 Analysis Approach

The analysis considered spectrophotometric measurements from three physical exemplifications of the Munsell system:

1. *Munsell Book of Color - Matte Finish Collection*, Munsell Color Corporation, 1976,
2. *Munsell Book of Color - Glossy Finish Collection*, X-Rite Corporation, 2007, and
3. Jim Long & Joy Turner Luke, *The New Munsell Student Color Set,* 2^{nd} *ed.,* Fairchild Publications, New York, 2001.

The measurements of most interest for this analysis were made by Finland's University of Joensuu in 1988 or 1989, of the 1976 *Munsell Book of Color*. The two more recent books serve mainly as comparison for Joensuu's data. They were measured by the author and an associate in 2012.

The question of interest is how well the physical exemplifications match the Munsell renotation. The Munsell renotation lists aimpoints, in CIE coordinates, for 2,745 standard Munsell colours. A Munsell colour is considered standard if it has an integer

Munsell value, an even Munsell chroma, and a hue prefixed with 0.0, 2.5, 5.0, 7.5, or 10.0. Non-standard colours are interpolations between the standard colours. All the books contain both standard and non-standard Munsell colours. For example, the first two books contain many colour samples with chroma 1; since 1 is not an even number, these samples are non-standard. Aimpoints for non-standard colours can be interpolated from standard aimpoints, but different investigators might interpolate differently. To avoid this source of variability, non-standard colours were discarded, and only measurements of standard colours were analyzed.

The spectrophotometric measurement of each standard colour produced a reflectance spectrum. The reflectance spectrum in each case took the form of reflectance percentages at a set of evenly spaced wavelengths. The limits of the wavelength, and the interval between adjacent wavelengths, varied with the spectrophotometer. Analysis was restricted to the visible spectrum, chosen to lie between 400 and 700 nm inclusive, so data outside this interval was not used. The Munsell renotation assumes that samples are viewed under Illuminant C, so the relative spectral power density (SPD) for Illuminant C, taken from Table II(3.3.4) of Reference 8, was multiplied by the reflectance percentage, at each wavelength. This product defines the spectral stimulus that a Munsell colour presents to a viewer's eye. Since the ambient illumination for the Munsell system is assumed to be Illuminant C, relative CIE coordinates[9] can be calculated, using the 1931 2° colour-matching functions.[10] The Munsell renotation[1] gives aimpoints for standard Munsell colours in relative CIE coordinates, so the printed samples can be compared with the aimpoints.

Comparisons were made in several ways. First, a histogram of ΔE_{00}s was produced, and summary statistics, such as the mean and median ΔE_{00} were calculated. This histogram gives an overall picture of the accuracy achieved. The CIEDE 2000 expression, as implemented in Ref. 5, was used to calculate ΔE_{00}s, and the three denominators in the expression were all set equal to 1. Second, biases in the printed samples were investigated. A bias would occur, for example, when colours were printed consistently too light, or when hues of printed colours tended more towards yellow than they should. A visual inspection of a Munsell book would likely overlook biases, because consistency would still be maintained. If an entire hue leaf were printed too light, for example, the value and chroma *relationships* could still be preserved, even though individual values and chromas were wrong.

The investigation of biases required inverting the Munsell renotation, that is, going from a measured set of CIE coordinates to a Munsell specification. Since the CIE coordinates were rarely a perfect match for any standard Munsell colour, a non-standard Munsell colour was found, using a recently published algorithm.[11] As an example, the CIE xyY coordinates of the measurement of the sample for 5B 7/8 were calculated to be

$$\begin{aligned} x &= 0.2211, \\ y &= 0.2692, \\ Y &= 35.27. \end{aligned} \tag{1}$$

The Munsell renotation coordinates for 5B 7/8 are

$$\begin{aligned} x &= 0.2204, \\ y &= 0.2729, \\ Y &= 41.99. \end{aligned} \tag{2}$$

The inverse renotation converts the xyY coordinates in Expression 1 into Munsell coordinates 5.7B 6.5/7.5. In comparison to 5B 7/8, we see that the printed sample is 0.5 value steps too dark, 0.5 chroma steps too dull, and 0.7 hue steps too purplish. These errors in the individual Munsell components allow a more refined look at the accuracy of printed books.

Tables 1 and 2 present the measurement details for the data used in the analysis, and some summary statistics.

Book	Printed	Measured	Spectrophotometer	Wavelengths (nm)
Munsell Book of Color—Matte Finish	1976	1988/9	AOTF	400-700, by 5
Munsell Book of Color—Matte Finish	1976	1988/9	Perkin-Elmer Lambda 9	380-800, by 1
New Munsell Student Color Set, 2nd ed.	2001	2012	ColorMunki Design	380-730, by 10
Munsell Book of Color— Glossy Finish	2007	2012	ColorMunki Design	380-730, by 10

Table 1: Characteristics of Munsell Book Measurements

Book	Number of Standard Colours	Median ΔE_{00}	Mean ΔE_{00}
1976 *Munsell Book of Color—Matte Finish*	1003	6.9	7.2
1976 *Munsell Book of Color—Matte Finish*	1021	3.6	3.8
New Munsell Student Color Set, 2nd ed.	235	1.6	1.8
2007 *Munsell Book of Color— Glossy Finish*	1301	1.9	2.1

Table 2: Summary Statistics for Munsell Book Measurements

3 Analysis Results

3.1 Joensuu Measurements of the 1976 Matte Munsell Book

As part of an extensive colour database,[6] researchers at Finland's University of Joensuu measured the 1976 matte edition of the *Munsell Book of Color*. Two sets of

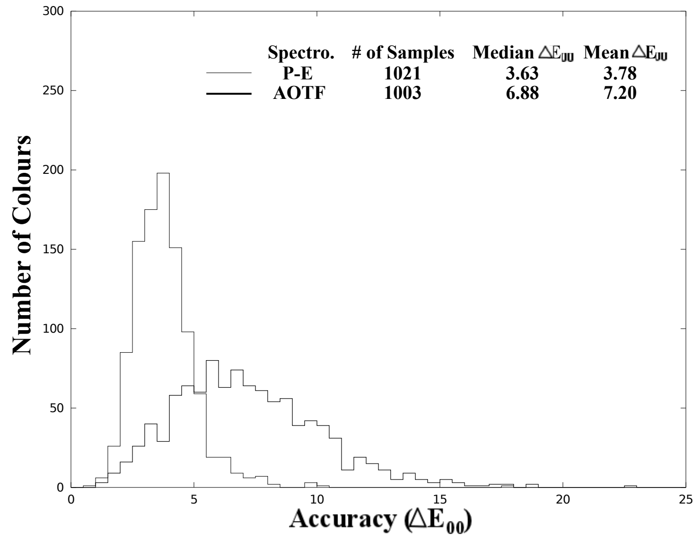

Figure 1: Accuracies for Joensuu Measurments of the 1976 *Munsell Book of Color—Matte Finish*

measurements were made,[7] one with an acousto-optical tunable filter (AOTF), and the other with a Perkin-Elmer Lambda 9 spectrophotometer. Both sets of reflectance spectra are publicly available.[7] The two sets differ significantly, as will be seen, and it is difficult to find any simple explanation.

The AOTF data was measured at wavelengths varying from 400 nm to 700 nm, in increments of 5 nm. The researchers noted that there was usually excess noise at 420 nm and below, so the analysis presented here discarded any reflectance data at wavelengths less than 425 nm. Since less than 1 percent of the colour-matching functions are below 425 nm, this data loss is inconsequential. The Perkin-Elmer data was finer, varying from 380 nm to 800 nm, in increments of 1 nm.

Figure 1 shows the ΔE_{00}s for both spectrophotometers. The Perkin-Elmer (P-E) spectrophotometer averaged ΔE_{00}s just under 4, while the AOTF spectrophotometer averaged ΔE_{00}s of about 7. Such ΔE_{00}s would be considered very poor matches for most applications. In an analysis similar to this one, Reference 12 calculated ΔE_{00}'s for some Joensuu measurements of the glossy 1976 Munsell book, rather than the matte book. The instrument for the glossy samples was a Perkin-Elmer Lambda 18 spectrophotometer, instead of a Lambda 9. Reference 12 found a mean ΔE_{00} of 3.5, which is not much different from our mean of 3.8.

Besides the overall colour difference, the individual Munsell components (hue, value, and chroma) were also analyzed, to identify possible biases. To investigate the

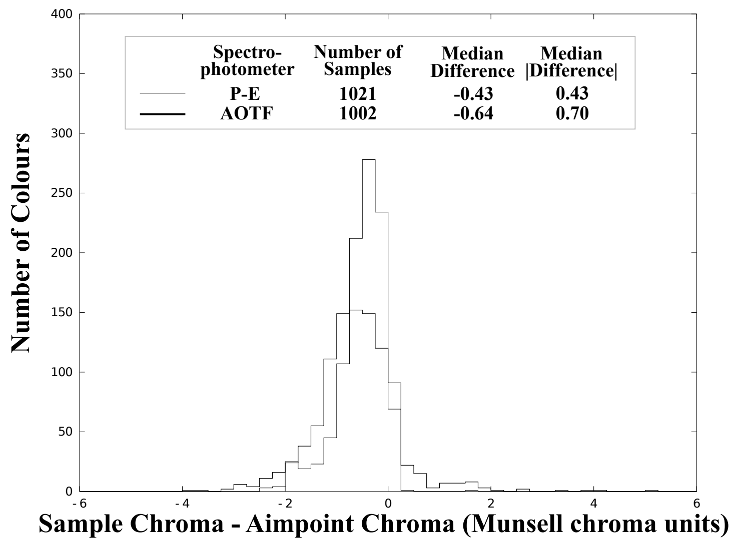

Figure 2: Chroma Shifts in Joensuu Measurements for the 1976 *Munsell Book of Color—Matte Finish*

individual components, the CIE coordinates of each sample were converted to the Munsell system, using Reference 11. Then the aimpoint's hue, value, or chroma, was subtracted from the sample's hue, value, or chroma, to measure the shift. A positive value shift means that the sample is lighter than the aimpoint, while a negative value shift means the sample is darker. Similarly, a chroma shift is positive when the sample is more saturated than the aimpoint, and negative when it is less saturated. By convention, hue shift is positive when red shifts toward yellow, yellow toward green, and so on around the hue circle; it is negative when green shifts toward yellow, yellow, toward red, and so on. When the aimpoint was a neutral colour, no hue shift was calculated. Figures 2 through 4 display shifts in the individual components.

Figure 2 shows that the chroma differences are almost all negative, indicating that the printed chromas are duller than the corresponding aimpoints. The median difference is about -0.4 chroma steps for the Perkin-Elmer spectrophotometer, and -0.6 chroma steps for the AOTF. Since the differences are mostly negative, the median absolute difference is about the same as the median difference. The AOTF measurements are slightly more dispersed than the Perkin-Elmer measurements, but on the whole the two sets agree well.

Figure 3 shows the hue differences for the two sets of measurements. The Perkin-Elmer measurements cluster closely near 0, indicating that there is little hue bias. The median Perkin-Elmer hue difference is just over 0.5 Munsell hue units. This

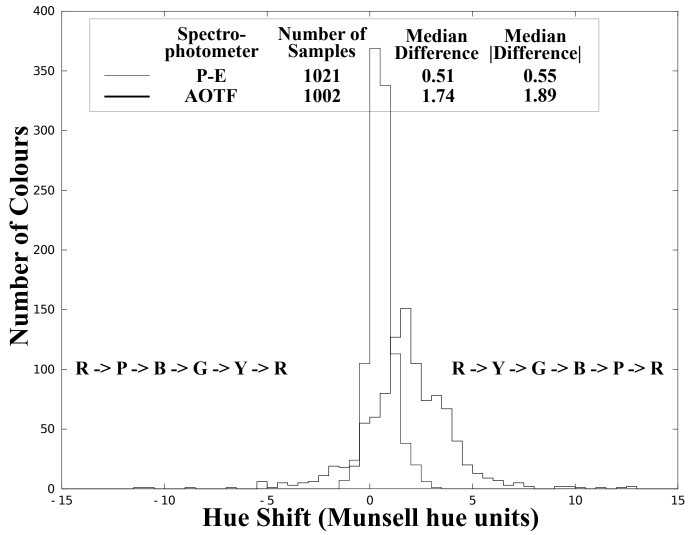

Figure 3: Hue Shifts in Joensuu Measurements for the 1976 *Munsell Book of Color—Matte Finish*

difference is barely perceptible, because even the standard difference of 2.5 hue units between adjacent hues is not perceptually large. The AOTF measurements, on the other hand, show a significant, consistently rightward shift of about 1.8 hue units. Many of these differences would definitely be noticeable. Furthermore, the AOTF differences are much more dispersed than the Perkin-Elmer differences. A similar difference in variability was seen in Figure 1. A simple interpretation is just that the AOTF spectrophotometer is less reliable than the Perkin-Elmer spectrophotometer.

Figure 4 shows that the value measurements disagree strongly in direction, though the median difference in each case is about 0.3 Munsell value units. The Perkin-Elmer value differences are almost all negative, indicating that the samples are printed darker than the aimpoints. The AOTF differences, by contrast, are almost all positive, indicating that the samples are printed lighter than the aimpoints. This striking discrepancy demands an explanation. Unfortunately, such an explanation is hard to come by.

Together, the AOTF results in Figures 2 through 4 suggest one reasonable explanation. The chroma shifts are predominantly negative, indicating that the printed colours are duller than their aimpoints. At the same time, the value shifts are predominantly positive, indicating that the printed colours are lighter than their aimpoints. The AOTF measurements also show a significant hue shift. A switch to lighter, duller colours of slightly different hues is consistent with fading: possibly the

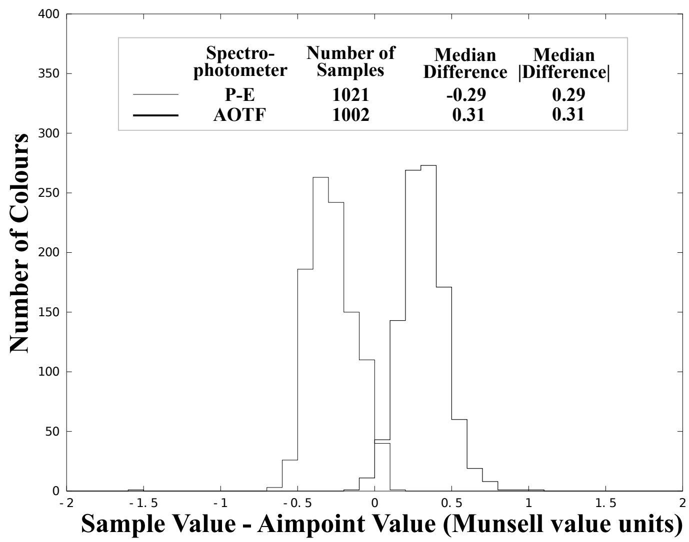

Figure 4: Value Shifts in Joensuu Measurements for the 1976 *Munsell Book of Color—Matte Finish*

colours were accurate when produced in 1976, but had faded by the time they were measured. The measurement date is uncertain, but cannot be later than 1989, when a paper[13] refers to them. A gap of a dozen years between printing and measurement might have led to fading, producing the consistent shifts and high ΔE_{00}s.

While this explanation is reasonable, it does not explain the Perkin-Elmer measurements, which were made after the AOTF measurements. The Perkin-Elmer measurements show a similar chroma shift to the AOTF measurements, and a hue shift in the same direction (though of about one-third the magnitude). The value shift, however, is in the opposite direction. As Figure 4 shows, the Perkin-Elmer measurements are about 0.3 value steps *darker* than the aimpoints, while the AOTF measurements are 0.3 value steps *lighter*. It is hard to reconcile this contradiction. A possible explanation is that the Perkin-Elmer spectrophotometer was used in specular exclusion mode: if this mode excluded some light that the AOTF would have measured, then consistently darker values would be expected.

In general, it is hard to make sense of the Joensuu data, and even harder to draw any conclusions about the 1976 *Munsell Book of Color*. Not only spectrophotometer characteristics, but also temporal colour changes, could be affecting the measurements. While the Perkin-Elmer data seems preferable to the AOTF data, the Perkin-Elmer data still shows ΔE_{00}s that would not be acceptable in many cases. If the Perkin-Elmer measurements are accurate, then the Munsell book itself is inaccurate, in which case it should not be a basis for further analysis. This examination, then, inspires little confidence in conclusions drawn from the Joensuu data.

3.2 The New Munsell Student Color Set, 2nd ed.

This Munsell set consists of swatches that a student sorts himself, and glues onto Munsell hue leaves. This book is limited in its gamut. Only hues prefixed with 5.0 are presented, and the chromas stay near the neutral axis. Figure 5 is a histogram of the differences between the printed samples and their renotation aimpoints. Out of 235 printed colours, the median ΔE_{00} was 1.6, and the mean ΔE_{00} was 1.8. A very large ΔE_{00}, of 10.2, occurs for the colour N9, which is printed far too light.

Besides the overall colour difference, the individual Munsell components (hue, value, and chroma) were also analyzed, to identify possible biases. The median value shift was 0.01 Munsell value steps, indicating that the samples were not consistently lighter or darker than their aimpoints. Even though the signed shift is small, the absolute shift could be large, which would indicate high inaccuracy. In fact, the median absolute value shift was only 0.05, which is well below human perceptibility. The median chroma shift was -0.02, and the median absolute chroma shift was 0.11, indicating likewise that chroma shifts are unbiased, and usually imperceptible. The median hue shift was -0.24, and the median absolute hue shift was 0.38. The distance between two adjacent standard hues is 2.5 Munsell hue units, so the hue shifts are also unbiased and minor. As a general conclusion, then, the errors in *The New Munsell Student Color Set* are unbiased in all three Munsell components, and no component is particularly inaccurate.

3.3 The 2007 Munsell Book of Color— Glossy Finish

For many years, X-Rite has produced its “Big Book of Color,” a Munsell book that is an industry standard. An extensive set of inks is used, to attain a wide gamut. More recent versions have over 1600 removable printed samples. In the 2007 book, 1301 printed samples were standard Munsell colours. Figure 6 is a histogram of accuracies. Their median ΔE_{00}, when compared to renotation aimpoints, was 1.9, and the mean ΔE_{00} was 2.1. Overall, then, this Munsell book achieved a ΔE_{00} accuracy of about 2.0, not much different from *The New Munsell Student Color Set.*

To test for bias, statistics were calculated for the individual Munsell components. The median value shift was 0.10 Munsell value units, and the median absolute shift was 0.11 units. The median chroma shift was 0.18 units, and the median absolute shift was 0.19 units. The median hue shift was -0.12 units, and the median absolute

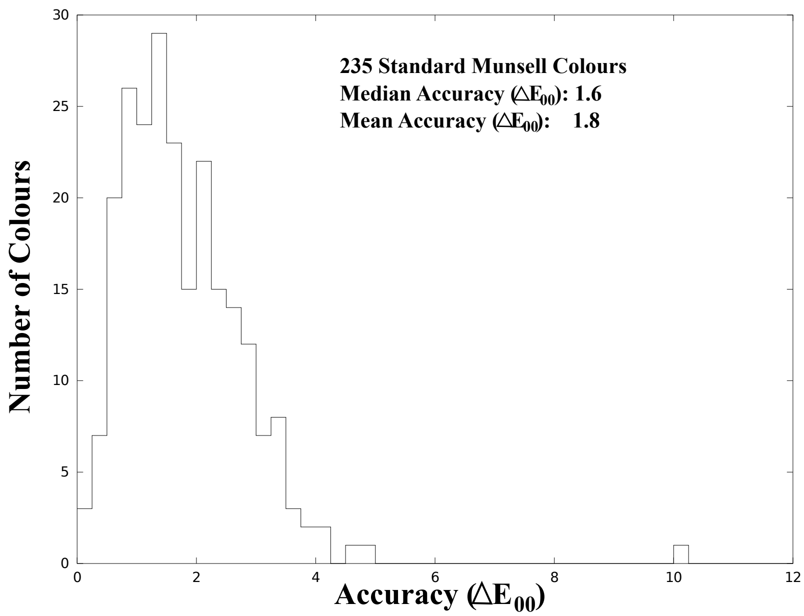

Figure 5: Accuracies for *The New Munsell Student Color Set, 2nd ed.*

hue shift was 0.34 units. Like *The New Munsell Student Color Set*, all these differences are near enough to 0 that we conclude that there is no consistent bias in any individual component, and no component is particularly inaccurate.

3.4 Comparison of Measurement Sets

Table 2 compares summary statistics for the Joensuu measurements of the 1976 Munsell book, with the more recent Munsell products. In addition, individual Munsell components have been analyzed for all measurement sets. In general, measurements of more recent Munsell books inspire more confidence. Not only are the recent differences smaller than the 1976 differences, but recent books also show no biases in hue, value, or chroma. The results obtained from the AOTF spectrophotometer, at least as employed for these measurements, are too variable to be of practical use. Since ΔE_{00}s near 7 are much too large for most colour applications, it is recommended that the AOTF measurements not be used. The Perkin-Elmer measurements, with average ΔE_{00}s just under 4, are preferable, but the two more recent sets, with average ΔE_{00}s just under 2, should be preferred to either of the Joensuu sets.

Apart from checking accuracy, reflectance spectra of Munsell books are helpful for further research, for example in dimensional analyses of colour.[13,14] A com-

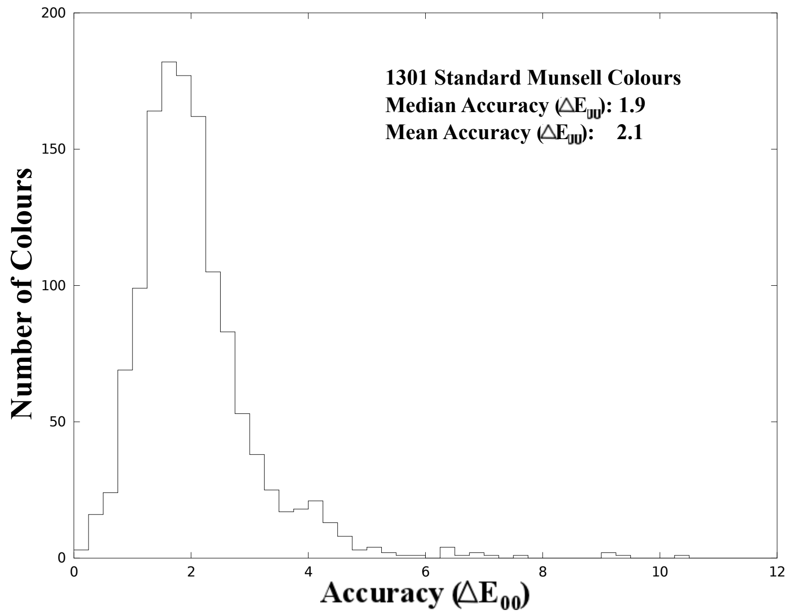

Figure 6: Accuracies for the 2007 *Munsell Book of Color— Glossy Finish*

mon approach is a principal components analysis (PCA), which tries to construct a small basis, such that every observed reflectance spectrum is a linear combination of basis elements. Mathematically, the set of possible reflectance spectra is infinite-dimensional, but empirical spectra (those that occur in actual surfaces) seem much more limited. Starting with physical samples, such as the swatches in a Munsell book, guarantees that the reflectance spectra are empirical.

Ideally, they should match the renotation aimpoints, too. One approach to achieve matching, taken by Derhak and Berns[12] for glossy samples, "corrects" the measurements. A correction modifies the measured spectra to make them match the aimpoints. If the modifications are small enough, then it is physically plausible that the corrected spectra can actually be produced. Since recent measurements have much smaller errors than the Joensuu measurements, small corrections would be more likely if one began with a recent sets of measurements, instead of the Joensuu data. As an overall conclusion, then, the lower ΔE_{00}s of more recent measurement sets make them a more suitable choice for research than the Joensuu measurements.

4 Summary

Four sets of Munsell measurements were considered:

1. AOTF measurements of the 1976 Munsell book,
2. Perkin-Elmer Lambda 9 measurements of the 1976 Munsell book,
3. ColorMunki measurements of the 2001 *New Munsell Student Color Set*, and
4. ColorMunki measurements of the 2007 Munsell book.

Because the ΔE_{00}s for the AOTF measurements were very large, with a median near 7, it is recommended that they be discarded. The Perkin-Elmer Lambda 9 ΔE_{00}s, with a median near 3.5, are an improvement, but still would not be considered accurate enough for many purposes. In addition, they show a consistent bias towards being too dark. Considerable correction would be needed to modify the measurements into physically realistic reflectance spectra that match the Munsell renotation. The two more recent sets of measurements have median ΔE_{00}s between 1.5 and 2.0. Their reflectance spectra agree with the Munsell renotation much better than the earlier sets, and would make a more accurate starting point for further analysis.

References

1. Sidney Newhall, Dorothy Nickerson, & Deane B. Judd. "Final Report of the O. S. A. Subcommittee on the Spacing of the Munsell Colors," JOSA, Vol. 33, Issue 7, 1943, pp. 385-418.
2. *Munsell Book of Color - Matte Finish Collection*, Munsell Color Corporation, Baltimore, Md., 1976.
3. *Munsell Book of Color - Glossy Finish Collection*, X-Rite, 2007.
4. Jim Long & Joy Turner Luke, *The New Munsell Student Color Set,* 2*nd ed.,* Fairchild Publications, New York, 2001.
5. Gaurav Sharma, Wencheng Wu, & Edul N. Dalal, "The CIEDE2000 Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations," COLOR Research and Application, Vol. 30, Number 1, pp. 21-30, February 2005.
6. Oili Kohonen, Jussi Parkkinen, & Timo Jääskeläinen, "Databases for Spectral Color Science," COLOR Research and Application, Vol. 31, Number 5, October 2006, pp. 381-390.
7. http://cs.joensuu.fi/ spectral/databases/download/munsell_aotf.htm, accessed October 30, 2012.

8. Günter Wyszecki & W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd ed.*, John Wiley & Sons, 1982.
9. Mark D. Fairchild. *Color Appearance Models, 2nd ed.,* John Wiley & Sons, Ltd, 2005, Sect. 3.6.
10. Deane B. Judd. "The 1931 I. C. I. Standard Observer and Coordinate System for Colorimetry," JOSA, Vol. 23, October 1933, pp. 359-374.
11. Paul Centore, "An Open-Source Inversion Algorithm for the Munsell Renotation," COLOR Research and Application, Vol. 37, Number 6, pp. 455-464, December 2012.
12. Maxim Derhak & Roy S. Berns, "Analysis and Correction of the Joensuu Munsell Glossy Spectral Database," presented at the 20th Color and Image Conference, 2012.
13. J. P. S. Parkkinen, J. Hallikainen, & T. Jääskeläinen, "Characteristic spectra of Munsell colors," JOSA A, Vol. 6, Issue 2, 1989, pp. 318-322.
14. Laurence T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," JOSA A, Vol. 3, Issue 10, 1986, pp. 1673-1683.

