Thresholds for Perception of Color Differences
The Challenge

The description of the InnoCentive Challenge entitled **Thresholds for Perception of Color Differences** included the following specifications:

"The seeker is open to creative solutions and concepts. However, solutions envisaged by the Seeker include:

The optimum solution would consider human visual sensitivity as a function of color/shade, customer/patient demographics and lighting. These would be correlated to the appropriate quantitative measurement parameters relative to color/shade and lighting. This would enable the seeker to integrate the subjective "human element" with a quantitative measurement resulting in a specification that would minimize product complaints for pharmaceutical product.

The Seeker wishes to better understand the following aspects relevant to setting limits for product color differences:

Methods to correlate quantitative color measurements with human perception of color differences that would result in a product complaint.

Best practices in managing visual perception focus groups and assessing focus group data.

Understanding population differences relevant to perception of color."

$5,000 Prize Solution for the Challenge

By A. Kimball Romney

Measuring perceptual color differences

Introduction.

The proposed solution for better understanding the perception of color differences and how they correlate with physical measurements (surface reflectance spectra) will be presented in three sections. First, a measurement model (consisting of a mathematical equation) for estimating where a human observer would locate any surface reflectance spectrum in Munsell conceptual space (given specified viewing conditions). Second, some suggestions about data collection and analysis for setting limits for product color differences based on consumer evaluations. Third, a brief review of what is known about population differences in color perception based on characteristics like gender, age, and ethnic or cultural background.

Measurement model.

Virtually all modern color specification systems begin with the definition of a Standard Observer. Our proposed measurement model uses CIE 10 degree 1964 functions shown in Figure 1 (left panel), available at http://cvrl.ioo.uc;.ac.uk/. The model uses D65 illuminant from the same site normalized to a maximum value of 1 as shown in Figure 1 (right panel).

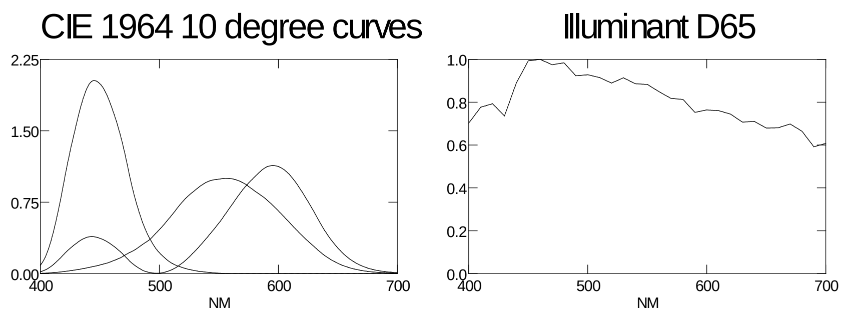

Figure 1. Curves from CIE 1964 10 degree curves and Illuminant D65

The first three data elements of the model are identical to the standard CIE color systems used in industry (for example L*a*b*). The major innovative feature of our proposal is to use the Munsell color system to represent the best available uniform isotropic color appearance system. It has been thoroughly researched for over a century and has some very desirable features. One of the most important is that there is an invariant relationship between the physically measured reflectance spectra of a surface and the color appearance locations assigned to it in Munsell color space. The space is orthogonal and isotropic which means that a unit of change in any part of the space is the same as in any other part of the space. Thus MacAdam (1981) error ellipses become circles. Complementary colors are connected by straight lines that pass through the origin of the system. The combination of mixing two colors on a spinning disc is calculated as the convex combination of the two. Calculations between the space of reflectance spectra and the color appearance space are valid and convenient (Romney, 2008). The use of Munsell is so important that we devote a few pages to the history of the Munsell system in Appendix A.

The proposed model is based on the model presented in D'Andrade and Romney (2003) with three minor modifications. The first modification is to use the CIE 1964 10 degree curves rather than the Stockman and Sharpe cone sensitivity curves to represent our Standard Observer. The second modification is to provide a precise linear transformation matrix in place of their "opponent process weights". The third modification is to skip the standardization procedure. With these modifications the cartoon illustrating the D'Andrade and Romney model in their Figure 3 is valid for the proposed model. In everyday words we could say the model begins with a measured spectrum, adjusts the spectrum by multiplying it by an illuminant such as D65, and multiplies the result by the CIE 1964 10 degree curves resulting in three numbers representing product sums. These three numbers are cube rooted. The standardization and subsequent weighting process in the D'Andrade and Romney model is replaced in the proposed model by a linear transformation specified in a three by three matrix.

The proposed model may be illustrated by applying it to estimate the location of the reflectance spectra (measured a percent reflectance on a scale of 0 to1) of the 1269 color chips from the 1976 Munsell Matte Color Book measured in 1 nm resolution from 400 to 700 nm available from http://cs.joensuu.fi/. We use linear algebra to represent large sets of reflectance spectra rather than one spectrum at a time. The model is defined by Equation 1:

$$\hat{\mathbf{M}}_{1269\times 3} = (\mathbf{S}_{1269\times 301}\mathbf{I}_{301\times 301}\mathbf{Q}_{301\times 3})^{1/3}\mathbf{P}_{3\times 3} \qquad \text{[Equation 1]}$$

Mathematica code for Equation 1: $\mathbf{M} = \sqrt[3]{\mathbf{S}\cdot\mathbf{I}\cdot\mathbf{Q}}\cdot\mathbf{P}$

In Eq. 1 the M matrix represents estimates (represented by the hat symbol) of the coordinates for each of the color chips in Munsell conceptual space. The idealized conceptual coordinates as well as the estimated coordinates (obtained with the Mathematica code shown above) of the chips is shown in Figure 2 where the axes are designated as: Munsell Value axis, Munsell 5 Red versus 5 Blue Green axis, and Munsell 10 Yellow versus 10 Purple Blue axis. The **S** matrix represents the reflectance spectra measurements of the chips, the **I** matrix represents a diagonal matrix with the illuminant D65 on the diagonal, the **Q** matrix represents the CIE 10 degree 1964 functions, and the **P** matrix represents a three by three linear transformation matrix that transforms the CIE function representation into Munsell space. We calculate it as the following:

$$\mathbf{P}_{3\times 3} = \begin{pmatrix} -0.0062 & 24.1570 & -6.7343 \\ 2.2174 & -21.8237 & 15.8368 \\ -0.1310 & -1.8836 & -9.0638 \end{pmatrix} \qquad \text{[Equation 2]}$$

Note that estimates for any reflectance spectra whatsoever may be obtained by substituting them in place of matrix **S**; however, matrix **P** is valid only for the specific Standard Observer and Illuminant specified. If either matrix **I** or matrix **Q** were to be changed it would require the computation of a new matrix **P**.

It may be seen that there is a close correspondence between the top panels and the bottom panels in Figure 1. We interpret the differences in the two plots to a variety of errors in locating the correct conceptual locations and in implementing their manufacture with available technology. See Appendix A for additional comments on the Munsell system.

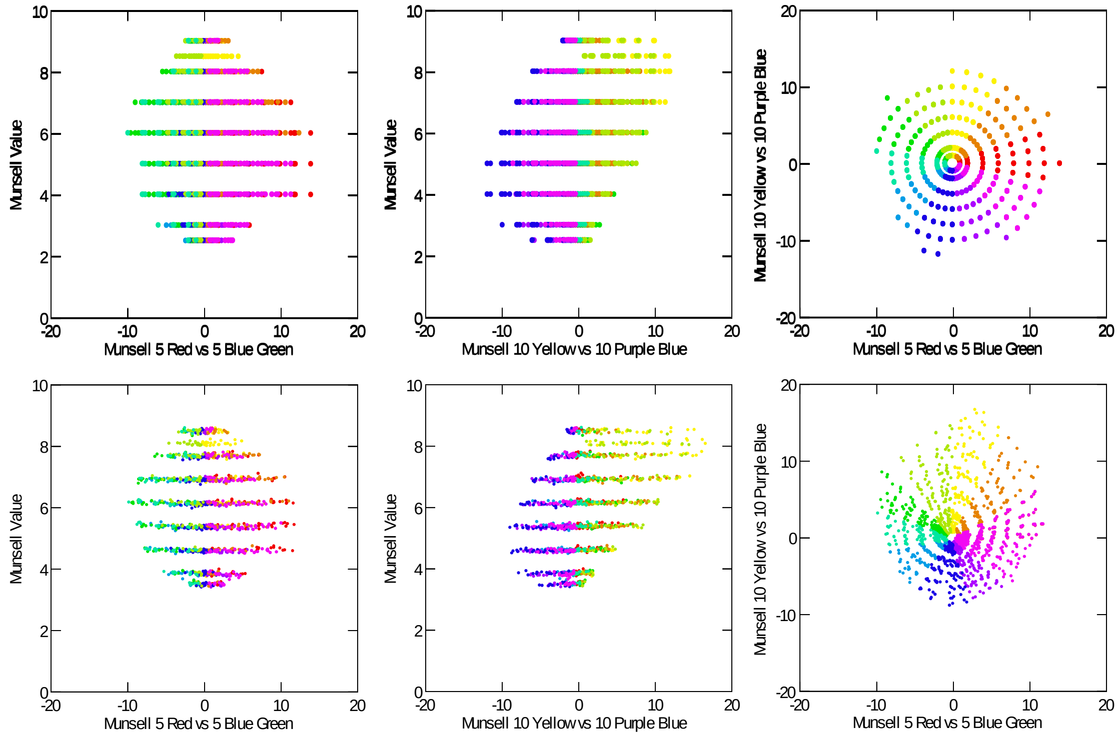

Figure 2. The three top panels show the Munsell conceptual locations of the Munsell Atlas chips plotted in Euclidean coordinates, while the three bottom panels show the estimated model coordinates of the measured reflectance spectra of the Munsell Atlas chips with the given conceptual labels.

It will be instructive to analyze an additional set of 108 reflectance spectra, both to help validate the Value axis and illustrate current state of the art accuracy (or inaccuracy). We include the 85 chips from the Farnsworth Munsell 100 Hue Test (test available from http://www.munsellstore.com/ for $688.75) to illustrate the extent to which they form a hue circle with chips of equal chroma and value. We include 21 neutral gray chips (Munsell value 3 through 8 in .25 steps) from a database of 1600 glossy Munsell color chips available from http://cs.joensuu.fi/ to verify that the model correctly aligns the vertical Munsell Value axis at the origin of the chromaticity plane. Finally we include two Planck blackbody radiation curves, one at 3900 K and one at 4000 K normalized to equal 0.5 at 545 nm. The pair of blackbody curves is included to help calibrate error circles in the model Munsell space since they appear as one of the highly investigated error ellipse pairs in MacAdam's (1981) famous study of just noticeable errors of color perception. The 108 reflectance spectra are shown in Figure 3.

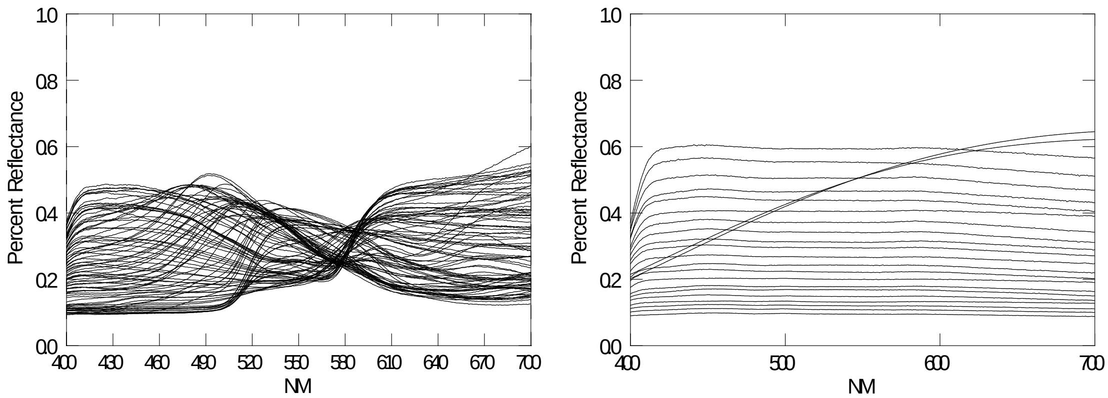

Figure 3. The reflectance spectra for the 85 Frarnsworth Munsell 100 Hue Test are shown in the left panel while reflectance spectra of 21 neutral gray chips and 2 blackbody spectra of 3900 K and 4000 K are shown in the right panel.

The Munsell color appearance of these 108 reflectance spectra may be estimated by designating them as matrix **S** in Equation 1. The results are shown plotted on top of the Munsell conceptual system in Figure 4 and in an enlarged version in Figure 5.

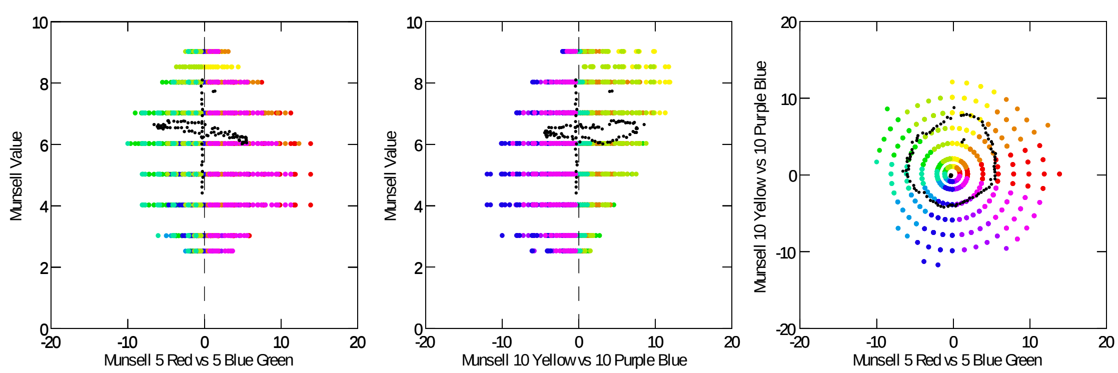

Figure 4. The estimated Munsell locations of the set of 108 reflectance spectra plotted as filled black dots on top of the Munsell conceptual system.

Several instructive features may be noted. First, the neutral gray chips are very closely aligned to the Munsell Value axis showing that the model is correctly oriented with respect to the three orthogonal Euclidean axes of the Munsell conceptual system. Second, the Farnsworth Munsell 100 Hue Test chips form a crude approximation of a circle varying from 6 to almost 7 on the Munsell Value axis and between 4 and 8 on the Munsell Chroma scale. The Planck blackbody spectra have higher Munsell Value and lower Munsell Chroma locations than the Farnsworth Munsell chips. A more detailed examination of the errors is shown in Figure 5.

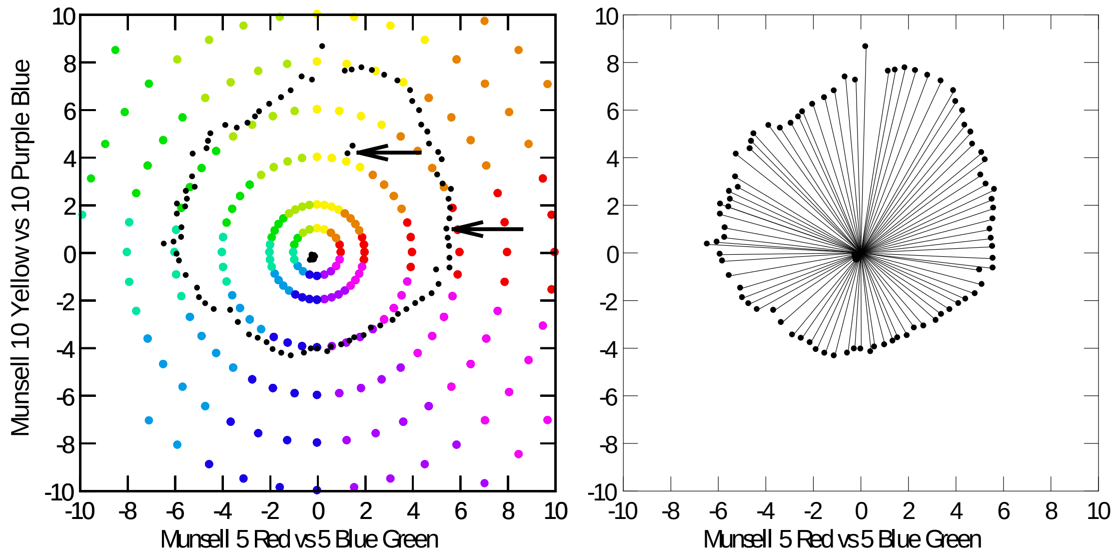

Figure 5. Two enlarged views of the Farnsworth Munsell chips. The upper arrow in the left plot points to the difference in the 3900 K and the 4000 K blackbody locations, an early standard of a fairly reliable just distinguishable color difference. The lower arrow points to the first chip in the Farnsworth Munsell color test which has the appearance of red. The plot on the right shows a vector plot which is a visual aid to examine the differences in the spacing of the hues in the Farnsworth Munsell color test.

Notice in the left panel of Figure 5 that the first chip in the Farnsworth Munsell series is at Munsell Chroma 6. The sequence proceeds counter clockwise up to chroma=8 in the yellow hues, back to chroma=6 in the blue green area, down to chroma=4 in the purple area, and finally back up to chroma=8 at the end of the sequence. Note that the 21 neutral gray chips all plot on top of each other in an area with a radius of 0.18 units, smaller than the difference between the two blackbody curves, which is 0.38.

MacAdam (1981, p. 136) defined a jnd (just noticeable difference) as twice the standard deviation of chromaticity matching by his subject (PGN). He notes that “a difference of 100 K of color temperature near 4000 K is slightly more than twice the standard deviation of color matching by PGN. Therefore, it is just noticeable”. The measured difference between the two blackbody curves in Figure 5 is 0.38 units in the Munsell conceptual space. If we compute the circumference of the chroma circles at different sizes we can check whether the spacing of the hues in the Farnsworth Munsell test are consistent with the blackbody results. In round numbers the circumference of chroma circles of 4, 6, and 8 are 25.13, 37.70, and 50.27 Munsell units. Dividing these numbers by 0.38 we get estimates of 66, 99, and 132, which represent the number of distinguishable hues (using MacAdam criteria) at those chroma levels. We conclude that there is a rough expectation that at Munsell Chroma=6 that the Standard Observer should be able to distinguish about 100 distinct hues. The test would be a lot more difficult at chroma=4 and a lot easier at chroma=8. This excursion into historic measures of error bounds of color perceptual differences may provide a benchmark that one would expect to exceed.

One purpose for examining this particular set of 108 reflectance spectra is to draw attention to the size of errors that characterizes some color products based on current color measurement practices. The example underscores the need to discover and to understand the source of these errors in order to correct for them. The work of MacAdam on error ellipses and the invention of the Farnsworth Munsell Hue Test were based on basic CIE methods that remain basically similar today. The CIE tristimulus values are valid but the transformations of that data lead to a misleading representation in the Chromaticity Diagram, which is not meant to represent color appearance and is neither orthogonal nor isotropic.

The proposed measurement model is based on the Standard Observer defined by CIE in 1964 and utilizes information identical to that used in the ordinary CIE tristimulus values but transforms that data to an orthogonal isotropic Euclidean color appearance system that has been well studied for over a century. The Munsell conceptual space we propose is intended to apply to comparisons of differences between object color samples (specified by measured reflectance spectra) of the same size and shape, viewed in an environment of a neutral gray of the same Munsell Value as the sample, viewed by an observer photopically adapted (to eliminate all rod activity) to a single field illuminated by adequate D65 light. **IMPORTANT CAVEAT**: Distances among chips in a single Munsell Value plane are valid when computed using Euclidean distance measures. Computation of Euclidean distances in the three dimensional space are not valid without adjustments to the Munsell Value scale (in the Munsell color system as used in the current atlases the scale of the units of Value do not correspond to the scale of the units of Chroma). Valid three dimensional distances would require a change in matrix **P**, which we could easily compute. Ordinary color difference judgments nearly equal on the value scale can ignore this adjustment.

A final important computational advantage of the model should be mentioned. The combination of reflectance spectra in the physical space of reflectance spectra follows the rule of convex combination. In the Munsell Euclidean model provided by Eq. 1 the rule for combining any two locations in the space follow the rule of convex combination. There is a biunique relation between the two sets of calculations that is well described in Romney (2008). This relation provides a powerful tool to go between measurements made by instruments and judgments of appearance by human subjects.

Suggestions about data collection.

Data collection is of critical importance since it is at this stage that many problems occur. The most important thing to keep in mind is that controlling all variables involved is crucial. Color contrast has huge effects so the context of viewing must be carefully controlled to avoid contrasting colors appearing in the same field of view. Even an observer wearing colored garments can affect the appearance of a sample. We suggest that the total testing environment be a neutral matte gray near Munsell Value=6 including the wall, floor, and furniture in the observation room. A neutral gray smock or outer garment should be provided for both observer and experimenter. The stimuli should consist of real material surfaces of actual and simulated products arranged in triads as in Figure 6 and illuminated by as near D65 standard as possible at an intensity well into the photopic range. Each stimulus set could consist of three samples for comparison mounted in the middle of a standard letter sized card stock of neutral gray (matched to be as near as possible the same Munsell Value as the samples). Figure 6 assumes three round pills each 1 cm in diameter separated by ½ cm. If the samples are some other shape they should be arranged to form a symmetrical pattern with the three items each equidistant from each other. If possible, samples should be flush with the surface of the background to avoid shadows (which affects the perception of color). Samples should also be free of indentations, such as logos or grooves to divide pills, since they may also affect perceptions.

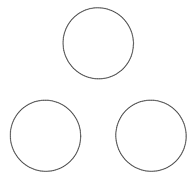

Figure 6. Suggested arrangement of stimuli to be compared.

The subject would be instructed to respond with one of the following responses: 1. all the colors are the same, 2. one of the colors is different from the other two (in which case the subject would indicate which is different with the alternatives, left, right, or top), and 3. all the colors are different from each other (in which case the subject would indicate which was most different from the other two with the alternatives, left, right, or top). Thoughtful designs for combining various stimuli (same batch versus different batches, damp versus dry samples, dark stored product with sun exposed product, etc.) will provide information for calibrating quality control.

We suggest that Seeker experiment with series of colors near to each other (for example, from low chroma blue to high chroma blue or transition among hues keeping value and chroma equal) to obtain guidelines as to what size differences subjects can actually detect. Under these observation conditions we believe one should be able to do considerably better than the 0.38 difference of the blackbody curves as well as radically better than the Farnsworth Munsell 100 Hue Test chips. When experimenting with the differences it is best to begin with changes along only one dimension at a time such as hue or chroma.

If the interest of Seeker is in the measurement of color perception then we would not recommend focus groups in which the subjects can influence other subjects by suggestions, leading remarks, force of personality, or other social factors. Data should be collected from one subject at a time free from outside influence. Aggregating independently obtained judgments generally provides more accurate results than those arrived at in face to face interaction. If group discussion is useful in other respects it should take place after data collection.

Seeker will want to conduct several designs to insure that adequate knowledge is obtained for quality control in realistic situations. The number of subjects required for each design need not be particularly large. In our experience 20 to 30 subjects should be adequate.

Population differences relevant to perception of color.

A recent study by Moore, *et al*. (2002) investigated cultural, gender, and individual differences in perceptual structures of colors in Chinese and English. They studied eight basic colors widely dispersed in chroma and hue but we believe their results apply to the judgment of small color differences as well. They studied 68 Chinese speaking subjects in Taipei, Taiwan and 52 English speaking subjects in the United States. Each subject completed three tasks involving color samples or color names. They used sophisticated quantitative measures to analyze the data and concluded: "The major findings are: (1) all respondents share approximately sixty percent of their knowledge of the judged similarity structures of both semantic and perceptual tasks, (2) there are genuine individual differences among respondents that account for about fourteen percent of their knowledge on average, (3) there are small but statistically significant gender differences that come to about three percent on average, (4) there are small but statistically significant differences between Chinese and English respondents of about one-and-a-half percent, (5) there are differences in the semantic structure of the names of colors as compared to the judgments of the color samples that amounts to about five percent, and (6) there is about a three percent difference in the paired comparison task and the triads test." (Moore, *et al*., 2002, abstract)

We believe that (considering the large error variance of their task compared to the experimental task outlined above) the main practical implication of these findings is that individual differences within the population groups (cultural/linguistic, gender) are several times as large as the differences between the population groups. In the controlled conditions specified above for data collection we believe that normal observers from any population (regardless of language, country of origin, or gender) will respond fairly similarly. Normal is defined as passing the Ishihara Color Test. Within any population group there will still be some individual variation related to factors such as genetic variability in the location of peak sensitivities in the photoreceptors. It is probably not practical to eliminate such variability even though in theory each individual could be modeled by obtaining a matrix **Q** specific for the individual (a color matching experiment or DNA could provide estimates of values to use) along with an appropriate matrix **P**.

If color blindness is an issue we might note that the most common types (protanope and deuternope) are deficient in distinguishing colors on the 5 Red versus 5 Blue Green axis. This suggests that if one were designing products easy for dichromats to distinguish, the colors should vary along the 10 Yellow versus 10 Purple Blue axis.

We might note that even though different populations may judge the differences among closely related colors in very similar ways, they may have very different preferences for colors. This suggests that if one were designing generic products, for example, where the customer exerts choices among products based in part on color, different populations might have very different preferences. This proposal has only considered judgments of perceived color differences without including preferences. We might note that there is a very large body of research on color preferences in different populations. The Japanese color literature is especially rich in such studies. There are also color taboos that occur in a variety of cultures.

References.

Note: We assume that Seeker is acquainted with standard works on color vision such as Wyszecki, G. and Stiles, W. S. (1982). *Color Science: Concepts and Methods, Quantitative Data and Formulae* (Wiley, New York), 2nd Ed.; Billmeyer, F. W. and Saltzman, M. (1981). *Principles of Color Technology* (Wiley, New York), 2nd Ed.; and Hunt, R. W. G. (2006). *The Reproduction of Color* (Wiley, New York), 6th Ed. We recommend that Seeker check the references below that are referred to in the proposal.

D'Andrade, R. G. and Romney, A. K. (2003). A quantitative model for transforming reflectance spectra into the Munsell color space using cone sensitivity functions and opponent process weights. *Proceeding of the National Academy of Sciences USA,* 100:6281–6286.

MacAdam, D. L. (1981). *Color Measurement: Theme and Variation.* (Springer-Verlag, Berlin).

Moore, C. C., Romney, A. K., and Hsia, T. (2002). Cultural, gender, and individual differences in perceptual and semantic structures of basic color in Chinese and English. *Journal of Cognition and Culture*, 2:101–134.

Romney, A. K. (2008). Relating reflectance spectra space to Munsell color appearance space. *Journal of the Optical Society of America A*, 25:658–666.

APPENDIX A. Brief summary of the Munsell color system.

The Munsell color system was developed in the early years of the last century by Albert H. Munsell. In many current accounts Munsell is described as an artist who devised a color system. In fact he did extensive scientific experiments and published scientific papers [3, 4]. He published a book [5] describing a color system together with an atlas. The color system uses three independent and orthogonal dimensions to represent all possible colors (as measured by reflectance spectra) in a distorted oblate shaped space, as illustrated in Figure 1.

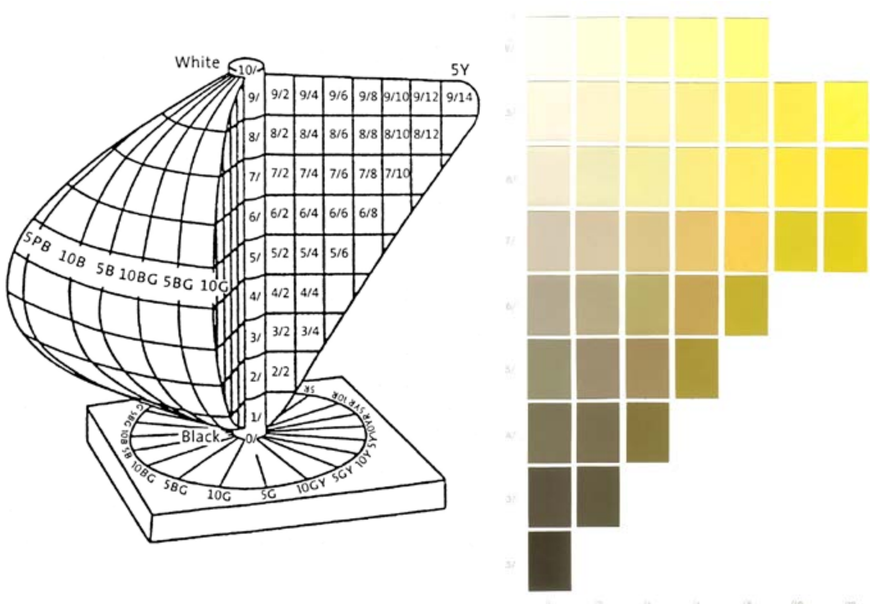

Figure 1. A view of the Munsell color solid and a sample page from the Atlas.

We cannot improve on Munsell's [3] own description, which "depends upon the recognition of three dimensions: *value, hue* and *chroma*. These three dimensions are arranged as follows. A central vertical axis represents changes in value from white at the top to black at the bottom….The value of every point on this axis determines the level of every possible color of equal value. Radial planes leading from this axis correspond each to a particular hue [e. g., the 5 Yellow Atlas page shown in Figure 1]. Opposite radial planes correspond invariably to complementary colors: any three planes separated by 120° form a complementary trio, etc. Thus the angular position of any hue is determined, and the hues are balanced. Chroma, or intensity of hue, is measured by the perpendicular distance from any point on the vertical axis, and the progression of chroma is an arithmetical one….Thus is constructed a solid. In this solid every horizontal plane corresponds to one and only one value. Every vertical plane extending radially from the central vertical axis contains but one hue. Finally, the surface of every vertical cylinder having the vertical axis as its principal axis contains colors of equal chroma."

Recent accounts of Munsell's color atlas usually fail to report on the extensive empirical research that he carried out using instruments that he invented (U.S. Patents 640,792, 1900; 686,872, 1901, 717,596, 1903, and 824,374, 1906), including a photometer and spinning top (Maxwell disk). Maxwell [[6]] and Helmholtz [[7]] did extensive experiments with spinning disks and knew that complementary colored surfaces would fuse to produce the appearance of an achromatic gray when spun at high speeds. This was seen as analogous to the observation that complementary monochromatic lights when mixed produce the appearance of white. Rood [[8]] presented extensive experimental data resulting from his research with spinning disks in his 1879 book. Munsell was greatly influenced by his study of Rood and followed his example by calibrating his color atlas with extensive experimentation. His complementary colors were all tested to produce an achromatic gray when combined with his spinning top. His equal chroma cylinders were produced by insuring that equal amounts (areas) of complementary colors produced achromatic grays. His equal spacing of hues was tested with triads of hues at 120° as well as complementary pairs. In addition he would take one of two complementary hues and balance it against a pair of hues separated by 72° to achieve an achromatic gray. In this way he could iterate to equally spaced hues as well as to calibrated value and chroma scales.

Munsell was aware of many complicating factors affecting color vision, including context effects such as color contrast, color induction, changes in illumination, etc. For example, in his first scientific publication [[4]] he was seeking to obtain the appearance of equally spaced chroma levels and found that he could only do so by specifying the observational context. He prepared two observational templates for a Maxwell disc as illustrated in Figure 2. Disk A is cut to produce a decreasing geometric progression of angles: 180, 90, 45, and so on, while the corresponding radii decrease arithmetically.

When a color (red in the Figure) is placed on a white background and spun, a well graduated progression of concentric rings is produced, diminishing in chroma and increasing in value from the center to the circumference. A change that prevents any variation in value may be obtained by changing the background to a neutral gray of the same value as the color; in which case the spacing of the chroma rings becomes very uneven. Disk B is now cut at five equal angles and with (as for disk A) an arithmetical decrease of radii. Now when a color is placed on a gray background and spun an even perceptual spacing of the chroma rings is observed. The finding is incorporated into the Munsell color system and requires colors to be judged on a neutral background (an achromatic gray of the same value). This reminds us of a very important general scientific principle, frequently overlooked in color experiments, which is to always isolate the variable (chroma in this case) being measured from possible interaction or confounding effect with other variables (value in this instance).

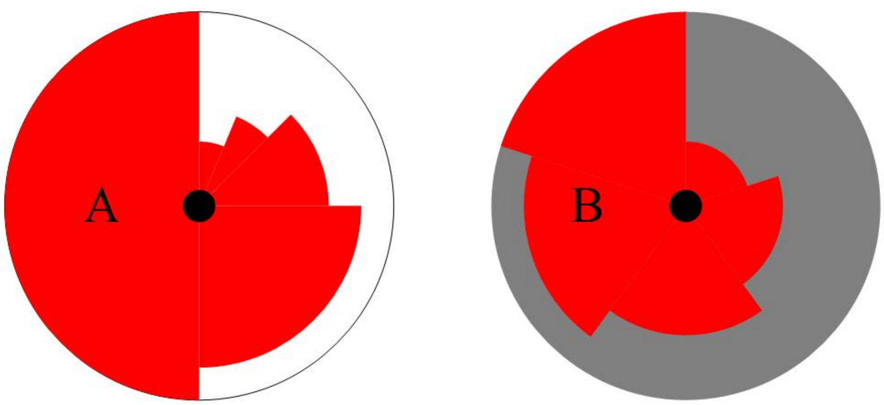

Figure 2. Munsell's two Maxwell disk templates.

When he produced his atlas, Munsell advised using as few pigments as possible and grinding them finely and consistently. Later, when the Optical Society of America [9] studied the Munsell color system, they produced samples exemplifying the Munsell renotation, all sample color chips being painted with only six colors plus black and white [10]. For reasons that are historically obscure the studies carried out by the Optical Society of America [11-15] did not use Munsell's psychophysical methods and substituted psychological judgments averaged over many subjects for their renotation system. We find only one replication of Munsell's careful experiments. It was written in 1940 by

Tyler (an undergraduate at MIT at the time) and Hardy [16] (who directed the preparation of the classic Handbook of Colorimetry [17]). They comment that their work "calls belated attention to the remarkable scientific insight of Professor A. H. Munsell. At a time when there was little to suggest such a procedure, he formulated rules for the construction of a psychophysical color system that could be used today without apology. We believe also that the publication of this paper may call attention to the fact that the psychophysical definitions of the terms hue, value, and chroma given in the Atlas of the original Munsell Color System differ from purely psychological definitions used since the death of Professor Munsell."

In the Munsell color structure, circles of equal chroma, but several different values, form cylinders. Any valid model of perceptual color space should reflect this fact. Research on scaling the similarity structure of the reflectance spectra of the Munsell atlas chips has usually reported a conical rather than a cylindrical structure for the chroma circles of different value; examples include Romney and Indow [18], Lenz, *et al.* [19], Koenderink, *et al.* [20], and Burns *et al.* [21]. We view these earlier models as describing the physical stimuli correctly but as misleading when interpreted as perceptual space. The cylindrical structure only emerges when a cube root transformation is applied to the product of reflectance spectra, the illuminant, and CIE function (or cone sensitivity curves) as in Eq. 1. We follow the D'Andrade and Romney (2003) procedure of cube rooting this product, since this is a critical element of their model and has the effect of transforming the physical space of a cone to the perceptual space of a cylinder. The use of the cube root transformation goes back to Plateau [22] and Stevens [23]. It is also used in the CIE L*a*b* colorimetric system [24].

REFERENCES

1. A. K. Romney, "Relating reflectance spectra space to Munsell color appearance space," J. Opt. Soc. Am. A **25**(3), 658-666 (2008).
2. A. K. Romney and C.-C. Chiao, "Functional computational model for optimal color coding," P. Natl. Acad. Sci. USA **106**(25), 10376-10381 (2009).
3. A. H. Munsell, "A pigment color system and notation," Am. J. Psych. **23**(2), 236-244 (1912).
4. A. H. Munsell, "On the relation of the intensity of chromatic stimulus (physical saturation) to chromatic sensation," Psych. Bull. **6**(7), 238-239 (1909).

5. A. H. Munsell, *A Color Notation*, 2nd Edition ed. (George H. Ellis Co., Boston, 1907).
6. J. C. Maxwell, "Experiments on colour, as perceived by the eye, with remarks on colour-blindness," Transactions of the Royal Society of Edinburgh **21**(2), 275-298 (1855).
7. H. von Helmholtz, "Physiological Optics," In Sources of Color Science, 84-100 (1866).
8. O. G. Rood, *Modern Chromatics* (Appleton and Company, New York, 1879).
9. K. L. Kelly, K. S. Gibson, and D. Nickerson, "Tristimulus specification of the Munsell Book of Color from spectrophotometric measurements," J. Opt. Soc. Am. A **33**(7), 355-376 (1943).
10. H. R. Davidson and H. Hemmendinger, "Colorimetric calibration of colorant systems," J. Opt. Soc. Am. A **45**, 216-219 (1955).
11. D. Nickerson and S. M. Newhall, "A Psychological Color Solid," J. Opt. Soc. Am. **33**(7), 419-419 (1943).
12. D. B. Judd, "The Munsell Color System," J. Opt. Soc. Am. **30**(12), 574-574 (1940).
13. K. S. Gibson and D. Nickerson, "An Analysis of the Munsell Color System Based on Measurements Made in 1919 and 1926," J. Opt. Soc. Am. **30**(12), 591-608 (1940).
14. J. J. Glenn and J. T. Killian, "Trichromatic Analysis of the Munsell Book of Color," J. Opt. Soc. Am. **30**(12), 609-616 (1940).
15. S. M. Newhall, "Preliminary Report of the O.S.A. Subcommittee on the Spacing of the Munsell Colors," J. Opt. Soc. Am. **30**(12), 617-645 (1940).
16. J. E. Tyler and A. C. Hardy, "An Analysis of the Original Munsell Color System," J. Opt. Soc. Am. **30**(12), 587-590 (1940).
17. A. C. Hardy, *Handbook of Colorimetry* (Technology Press, Cambridge, MA, 1936).
18. A. K. Romney and T. Indow, "Munsell reflectance spectra represented in three dimensional Euclidean space," Color Res. Appl. **28**, 182-196 (2003).
19. R. Lenz and P. Meer, "Non-Euclidean structure of the spectral color space," Proceedings of SPIE **3826**, 101-112 (1999).
20. J. J. Koenderink, van Doorn, Andrea J., *The structure of colorimetry,* Algebraic Frames for Perception-Action Cycle (Springer, New York, 2000), pp. 69-77.
21. S. A. Burns, J. B. Cohen, and E. N. Kuznetsov, "The munsell color system in fundamental color space," Color Res. Appl. **15**(1), 29-51 (1990).
22. J. Laming and D. Laming, "J. Plateau: On the measurement of physical sensations and on the law which links the intensity of these sensations to the intensity of the source," Psychological Research **59**(2), 134-144 (1996).
23. S. S. Stevens, *Psychophysics: intruduction to its perceptual, neural, and social prospects* (Wiley, New York, 1975).
24. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. (John Wiley & Sons, 1982).