6-16-2000

Lightness Dependencies and the Effect of Texture on Suprathreshold Lightness Tolerances

Ethan D. Montag
Rochester Institute of Technology

Roy S. Berns
Rochester Institute of Technology

Recommended Citation

Montag, E. D. and Berns, R. S. (2000), Lightness dependencies and the effect of texture on suprathreshold lightness tolerances. Color Res. Appl., 25: 241–249. doi:10.1002/1520-6378(200008)25:4<241::AID-COL4>3.0.CO;2-E

Ethan D. Montag[1]*
Roy S. Berns[1]
[1]Munsell Color Science Laboratory
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
54 Lomb Memorial Drive,
Rochester, New York 14623-5604

*Correspondence to: Dr. E. Montag, Munsell Color Science Laboratory,
Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology
54 Lomb Memorial Drive, Rochester, New York 14623-5604

Lightness Dependencies and the Effect of Texture on Suprathreshold Lightness Tolerances

A psychophysical experiment was performed to determine the effects of lightness dependency on suprathreshold lightness tolerances. Using a pass/fail method of constant stimuli, lightness tolerance thresholds were measured using achromatic stimuli centered at CIELAB L = 10, 20, 40, 60, 80, and 90 using 44 observers. In addition to measuring tolerance thresholds for uniform samples, lightness tolerances were measured using stimuli with a simulated texture of thread wound on a card. A texture intermediate between the wound thread and the uniform stimuli was also used. A computer-controlled CRT was used to perform the experiments. Lightness tolerances were found to increase with increasing lightness of the test stimuli. For the uniform stimuli this effect was only evident at the higher lightness. For the textured stimuli, this trend was more evident throughout the whole lightness range. Texture had an effect of increasing the tolerance thresholds by a factor of almost 2 as compared to the uniform stimuli. The intermediate texture had tolerance thresholds that were between those of the uniform and full-textured stimuli. Transforming the results into a plot of threshold vs. intensity produced results that were* *more uniform across the three conditions. This may indicate that CIELAB is not the best space in which to model these effects.*

Key words: color differences, color tolerances, parametric effects

INTRODUCTION

The Munsell Color Science Laboratory has been actively engaged in color-science research since its inception in 1983. One aspect of this research has been the evaluation of visual color differences towards the goal of improved specification of instrumental tolerances for industrial color control. This research has included the development of a color-tolerance[†] dataset for use in the testing and development of color-difference equations and the development of methodology for collecting and evaluating color-difference data.[1-4] In 1995, the Munsell Industrial Color Difference Evaluation Consortium was established to improve specifically the effectiveness of automated industrial-color difference evaluation.

In recent years, equations such as CIE94[5] and CMC[6] have been adopted for use in industry. These equations differ from their predecessors (CIELAB, ANLAB, and HunterLab) in that they derive tolerance ellipsoids based on the rotated lightness, chroma, and hue coordinates as opposed to the rectangular coordinates of CIELAB space.[7] Both these equations have parameters to weight the relative importance of lightness, chroma, and hue in the total color difference as functions of the position of the colors in color space. Additional weights can be adjusted to compensate for parametric changes in the samples and viewing environment based on empirical experience.

[†] The term "tolerance" is used in the context of this experiment to represent the perceptibility of color differences in comparison to a standard color difference set by an anchor pair. Because we are measuring suprathreshold color differences in comparison to a standard anchor pair color difference, we maintain the use of the term "tolerance" to differentiate these measurements from measurements of difference limen.

One specific difference between the CIE94 and CMC equations is the S_L term used to modify the contribution of lightness to the total color difference. In the CMC equation the contribution of lightness changes as a hyperbolic function of L* so that as L* increases, the contribution of lightness to the total color difference decreases. This term is set to unity in the CIE94 equation so that the contribution of lightness to the total color difference is constant at all lightness levels. Berns[8] argued that the use of the S_L term in the CMC equation may be a result of parametric differences between the data sets used to derive the equations. That is, the lightness dependency seen in the CMC equation may be due to the textile samples used in the experiments on which the CMC equation was based.[9] Berns found that eliminating this function improved the fit of the CMC equation to the RIT-Dupont data which were based on smooth, glossy painted aluminum panels. It should be noted that texture was but one of many experimental factors that differed between the sets of data upon which the equations were derived.

The purpose of this experiment is to measure suprathreshold lightness tolerances using uniform and textured stimuli. In addition to adding to the data-base for use in developing and refining instrumental color tolerance equations, we explore whether the differences in lightness dependency between CIE94 and CMC is a result of parametric differences in the textures of the samples used to measure the color tolerances. In these experiments the color samples were presented on a computer-controlled CRT display. Based on our previous experiments,[4] we found that this method is feasible and produces results that are comparable with object-color results. In addition, the use of the CRT decreases uncertainty and provides greater economy in time and cost for preparing stimuli. With the use of the CRT we can simulate texture and produce stimuli of the desired color difference quickly and easily.

EXPERIMENTAL

The experimental procedure consisted of a pass-fail psychophysical technique in which observers compared test pairs with an anchor pair. The observers' task was to indicate whether the anchor pair or the test pair had the greater overall color difference. Forty-four color normal observers, ranging from 18 to 45 years of age, took part in the experiment. The observers had varying degrees of experience making this type of judgment ranging from complete novice to well-practiced expert.

All stimuli were presented on a Sony Trinitron Multiscan 15sf monitor which was modified by the manufacturer to increase its peak luminance. The monitor was controlled by a Macintosh PowerPC. The software used to set-up and run the experiments was written in MATLAB[10] using the extensions provided by the high-level Psychophysics Toolbox[11] and low-level VideoToolbox.[12] A Radius ThunderPower 30/1920 video card was used in conjunction with the software to enable 10-bit resolution per channel.

The monitor was calibrated and characterized using the Gain, Offset, and Gamma method with an additive term for ambient flare and interreflections as described in references 13-15. As described in Montag and Berns,[4] the analytic CRT model was used to generate digital RGB values for the target colors to be used in the experiment. However the actual colorimetric values of the stimuli presented on the screen were measured and these measured values were used in the data analysis.

Generating the Stimuli

Suprathreshold lightness tolerances were measured around neutral color centers centered at CIELAB L*=10, 20, 40, 60, 80 and 100. These tolerances were measured for simulated full-textured samples, uniform samples, and samples with a simulated texture intermediate between the two.

Simulating texture

A number of samples of sewing thread which were wound on cards, similar to the samples used to derive the CMC equation, were provided to us by M. Ronnier Luo. These samples were imaged using an IBM Pro/3000 Digital Camera System. Each sample was scanned at high resolution (3000 pixels by 4000 pixels) through three filters (RGB) to produce an image. The images were cropped leaving a central homogeneous textured region (500 pixels by 500 pixels) eliminating the image of the card on which the thread was wound and the edges where the thread wraps around to the back. This area measured approximately 5 cm by 5 cm on the samples. These images were then scaled to a size of 200 x 200 pixels which was the size of the images used in the experiment. The goal was to produce images that had an appearance that was characteristic of the appearance of the actual samples. Sample images are shown in Figure 1. (The names assigned to the images are based on labels attached to the back of samples; they are not indicative of the actual or desired color appearance of the samples or their images.)

The software used to control the video board limited the number of simultaneously displayed colors to 256. Therefore the number of colors in the images was reduced to 120 or fewer to allow the simultaneous display of two textured images (plus anchor pair colors, black, white and the background). The number of colors was reduced using a minimum variance quantization routine in MATLAB which in some cases led to a much smaller number of colors in the image (e.g., the "Red" image contained only 51 colors).

The four images shown in Figure 1 were chosen as templates to create the neutral textured samples used in the experiment. They were chosen due to their quality and their range of lightness from the samples provided. The digital images displayed on the CRT were considered the original images and the digital RGB values were converted into CIELAB L*, C^*_{ab} and h_{ab} values based on the monitor colorimetric characterization.

The image histograms for the four template images are shown in Figure 2. The L*, C^*_{ab}, and h_{ab} histograms are shown in panels a, b, and c, respectively. The "red" image, which had a mean L* value of 11.5, was used as the template for stimuli around the L* = 10 color center. (A darker sample of black thread did not image well with our digital camera set-up.) The "brown" image, which had a mean L* value of 23.9, was used as a template for the stimuli centered on L* = 20 and 40. The "green" image, which had a mean L* value of 58.9, was the template stimuli centered around L* = 60. The "barley" image, which had a mean L* value of 71.9, was the template for samples centered around L* = 80 and 90. This was the lightest thread sample in the set. Each template image was processed using the following procedure:[16]

1) A 3-by-n matrix, **D**, where n is the number of colors in the image, is produced by subtracting the mean L*, C^*_{ab}, and h_{ab} values from the 3-by-n matrix, **x**, which is the list of L*, C^*_{ab}, and h_{ab} values of the colors in the template image.

2) The covariance matrix **C** is calculated by $\mathbf{C} = \mathbf{D}\mathbf{D}^t$, where $\mathbf{D}^t$ is the transpose of **D**.

3) Using singular-value decomposition (SVD) calculate **U**, the orthonormal matrix, and the diagonal matrix $\mathbf{S}^2$. The SVD algorithm decomposes the covariance matrix into the product of three components, $\mathbf{C} = \mathbf{U}\mathbf{S}^2\mathbf{U}^t$.

4) The matrix $\mathbf{M} = \mathbf{S}^{-1}\mathbf{U}^t$ is computed where **S** is the diagonal matrix containing the square-root of the elements in **S**.

5) Next calculate $\mathbf{B} = \mathbf{M}\mathbf{D}$. **M** is a 3-by-3 which when multiplied by **D** produces a decorrelated color space, **B**, with three independent dimensions.

6) To create samples with the same texture as the original templates, the inverse of **M**, $\mathbf{M}^{-1}$, is used to create $\mathbf{D}_{new}$ by $\mathbf{D}_{new} = \mathbf{M}^{-1}\mathbf{B}$ and new mean L*, C^*_{ab}, and h_{ab} are added to give $\mathbf{x}_{new}$ with the appropriate mean values. The mean C^*_{ab} and h_{ab} values were set equal to zero creating neutral samples with the same texture as the templates. In this case the transformation with matrix **M** is unnecessary but was done so that all the samples were created in a consistent manner.

7) To create uniform samples with no texture, matrix **B** was set to all zeros and the appropriate means were added after the inverse transform with $\mathbf{M}^{-1}$ was calculated.

8) Intermediate textures were created by multiplying **B** by 0.5. After the inverse transform the desired means are added to create the target samples with $\mathbf{x}_{new}$ values.

After the samples target samples were generated, the mean L*, C^*_{ab}, and h_{ab} values for each sample image were calculated. In many instances there was a small drift in C^*_{ab} away from zero. The magnitude of this drift was determined and the samples were recreated to correct for it. Figure 3 shows examples of full-texture, half-texture, and no-texture images constructed for the experiment.

For each color center (L* = 10, 20, 40, 60, 80, 90), 101 samples were generated varying L* in ±5 units in steps of 0.1. This was done for each texture type (full texture, half-texture, and no texture). In this way 1818 images were generated.

On three successive evenings all the images, the background, white, and anchor pair colors were measured using an LMT C-1200 colorimeter which was interfaced to the computer for automated measurement. The images were tiled and cropped to fill the area of the detector head. The spectral radiance of the three monitor phosphors were measured at each channel's peak output. These spectra were used to convert the CIE 1931 2° readings from the colorimeter to CIE 1964 10° values. The average of three readings were used for sample selection.

For each color center and texture the ΔL^*, ΔC^*_{ab}, and ΔH^*_{ab} values for every possible pair were calculated. From this list sample pairs were selected so that the ΔL^* values spanned an appropriate range centered on the color center and ΔC^*_{ab} and ΔH^*_{ab} were minimized. The 10-bit per channel resolution allowed for the selection of samples so that the contribution of L* to the total color difference was at least 97.76 %.

A pilot experiment was run on five observers in order to adjust the range of ΔL^* for the experiment. Ten sample pairs were chosen for presentation for each color center and texture type. Therefore there were 180 trials (10 trials x 3 textures x 6 color centers) per observer. Observers completed the experiment in about one-half hour. During the course of the experiment it became apparent that for many observers, the tolerance thresholds for the textured stimuli were larger than those in the pilot experiment. For these series, additional sample pairs were added to extend the upper end of the range and sample pairs were removed from the lower end. Because of this, not all the samples pairs in a particular analysis have the same number of total presentations. This change occurred after 20 observers had participated so that in some cases a tolerance threshold is based on 20 judgements at the low end, 44 judgements in the middle, and 24 judgements at the high end. This did not lead to any problems in the analysis.

Stimulus Presentation

Because the textures are identical for all the samples for each color center, the program that displayed the stimuli made a copy of each image that was upside-down and reversed. Upon presenting a stimulus pair, one image was in the original orientation and one image was the inverted copy. The position of these was altered at random. In this way, observers were prevented from comparing identical features in the texture and could make their judgements on the overall color difference.

Figure 4 is a picture of the experimental display. The anchor pair and stimuli patches were 6.5 cm on a side subtending an angle of 6.2° at a viewing distance of 60 cm. A one-pixel wide black border surrounded each sample patch. Observers free-viewed the samples at a normal viewing distance for a CRT display. A 1 cm wide white border (close to D65) surrounded the display defining the reference white. The anchor pair, consisting of two uniform patches with no texture, had a $\Delta E^*_{ab} = 1.0$ divided in L*, a*, and b*. The colorimetric values of the display are shown in Table I.

The experiments were performed in a darkened room. Observers adapted for two minutes to the background and border display before beginning the session. The anchor pair was always presented on the left side of the display however the top and bottom position of the anchor pair was randomized for each trial. All the trials for the six color centers and three textures were intermixed and presented in a unique random order for each observer.

RESULTS AND DISCUSSION

The data were analyzed using probit analysis,[17,18] a univariate statistical method that locates a median threshold from binary choice (in this case, greater or less-than) visual data. An overall fit of the data is assessed by a chi-square test and a heterogeneity factor is applied if the probability of chi-square was less than 0.1. The heterogeneity factor was applied for the no-texture analyses for color centers located near L* = 10, 20, and 30. The T50 (50% tolerance level) and fiducial limits were calculated for each color center.

The results are presented in Figure 5 and Table II. The abscissa is the average L* values of the test samples used in each series. The ordinate is the value of the threshold tolerance. The 95% fiducial limits are plotted for each point. The arrow on the abscissa indicates the lightness of the background ($L^* = 51.0$) used in the experiment.

Lightness dependency is evident for all three textures. For the uniform samples, the lightness tolerances remain fairly constant up to $L^* = 60$ and then increase fairly rapidly from a value near 3 ΔL^* to nearly 6 ΔL^* units. There is no clear evidence of a crispening effect[19] for the uniform samples. Crispening is the term used to describe the phenomenon in which the perceived magnitude of a color difference is increased when the color of the two stimuli being compared are similar to the background. This would result in a dip around $L^* = 50$. It is possible that such an effect exists for the uniform samples but is not evident in this plot because the tolerance for samples centered near $L^* = 50$ was not measured.

For the full-texture samples, the tolerance threshold can be characterized as increasing throughout the range except for a dip at the $L^* = 40$ color center. This dip may be a result of crispening. If such an effect is present, it is much larger in the textured stimuli as compared to the uniform samples, as plotted in Figure 5. In addition it would seem to be asymmetrical as plotted as a function of L^* since the threshold at $L^* = 40$ is much lower than the one at $L^* = 60$. The results for the half-texture samples show the same trends as those of the full-texture samples except the magnitude of tolerances is about one L^* unit smaller.

Due to the large fiducial limit for the textured stimuli at the higher lightness levels, it is difficult to judge whether the increase in tolerance is continuous or whether it begins to level off at the higher lightness levels. The large uncertainty in these tolerances may be due to the range of samples used in the experiment. It is possible that a larger range of color-difference pairs may have produced results with more precision.

Figure 6 shows the results of the probit analysis for determining the threshold for the full-texture samples at L* = 80 where the fiducial limits are the largest. Analyzing the first 20 observers gave a determination of threshold at a ΔL* value of 7.05 when the range of test pairs extended from a ΔL* of 4.7 to 7.5 (Fig. 6a). For the second set of observers, the range extended from 5.6 to 8.4 ΔL* (Fig. 6b), with a T50 determined to be at 8.74. Combining the data from the two groups yields a threshold value of 9.26 (Fig. 6c). The complete test ranges for each texture and color center is presented in Table II.

Shifting the test range caused the resulting threshold to shift. This range effect is a common artifact in the method-of-constant-stimuli. The usual range effect causes the threshold to be located near the middle of the tested range. In these experiments, for the textured-stimuli, shifting the tested range to larger ΔL* values caused an even larger shift of the threshold values to higher values. It is clear from Figure 6 and Table II that in certain cases the T50 values are estimated by extrapolation in the probit analysis. However, the results from the probit analysis indicate that in these cases the Chi-Square goodness-of-fit tests of the probit model are all insignificant. The increased fiducial limits are therefore indicative of any uncertainty in T50 estimation.

Figure 7 is a plot of the lightness dependency functions from three color tolerance equations, CMC, the Leeds Color Difference equation (LCD),[20] and the BFD equation.[21] (CIELAB and CIE94 do not have lightness dependencies and would therefore plot as straight horizontal lines with a value of one.) All three of these functions underpredict the magnitude of color difference found in this experiment. In this experiment the range of T50 values varies from 3 to over 9 units of ΔL*. This may be due to the fact that the lightness of the anchor pair colors are close to the background lightness. The crispening effect may enhance the color difference in the anchor pair. More data is needed in order to explain this difference.

The LCD function is based on analysis of the RIT-Dupont data set,[2] the Luo and Rigg[22] data set and a data set collected by Kim.[20] The data in the Kim data set were collected using uniform painted samples. The LCD S_L function is set to unity at $L^* < 50$ and increases to a value of 2 at $L^* = 100$. It is similar to the data from the no-texture samples which demonstrates greater lightness dependency as L* increases and is relatively flat at $L^* < 50$. The BFD lightness dependency is due to a different scaling of lightness which is not based on L* but on a logarithmic function of luminance. There is no specific S_L term in BFD that weights ΔL as a function of increasing lightness.

If the tolerance thresholds are considered as suprathreshold increments (or decrements) between the two test stimuli equivalent to the magnitude of the difference between the anchor pair samples, we can replot the data as threshold versus intensity (TVI) curves (see, for example, ref. 23). To do this the mean L* value of the test samples for each color center is considered the background intensity (I_{L^*}) and the tolerance threshold is considered the increment (ΔI_{L^*}) for this background. Summing the two gives the total lightness for the incremental stimulus ($\Delta I_{L^*} + I_{L^*}$). The $\Delta I_{L^*} + I_{L^*}$ values and the I_{L^*} values from the experiment were converted to the corresponding Y tristimulus values for the CIE 1964 standard observer. (For example, the monitor white point has a Y tristimulus value of 154.) By subtraction, ΔY values were determined for each Y background. The same procedure was used to convert the 95% fiducial limits to luminance values.

Figure 8 shows the TVI plot of the data with log(Y) on the abscissa and log(ΔY) plotted on the ordinate. The arrow indicates the luminance of the background. In this plot the three curves are more similar in shape. Near the background luminance the curves show a dip which may be the result of crispening. These dips are more symmetrical around the background compared to the data in Figure 5. The dotted lines show the best fitted line to each curve using linear regression. These fitted lines have slopes which are less than one demonstrating sub-Weber Law performance which may be expected since these data are not true thresholds. These slopes are 0.75, 0.78, and 0.76 for the full-, half-, and no-texture samples, respectively.

Plotting the data in this way indicates that CIELAB space may not be the optimal space for the modeling of color tolerances. For differences on the order of 1 ΔE^*_{ab}, perceptual spacing may be better represented in a logarithmically scaled space as opposed to a space based on power functions such as CIELAB. It is possible that representing these differences in CIELAB may exaggerate small differences at higher L* values and disguise important features. More data is needed to resolve this issue.

To complete this analysis, a family of straight lines, with slopes equal to the mean of the slopes of the three fitted lines in Figure 8, was computed spanning the range from no texture to full texture by changing the y-intercept. The values of the points on these lines were then converted into L* values. The ΔL* vs. L* plots of these values are shown in Figure 9. Overlaid on this plot are the tolerance thresholds (repeated from Fig. 5).

The lines based on the straight line fits to the data plotted as TVI curves do not fit the experimental data well. However, based on the analysis in Fig. 8, it may be worthwhile to consider these curves as good first approximations for estimating lightness dependency and the effect of texture on lightness dependency without the influence of crispening. These curves are well fit by quadratic functions. Only the CMC equation has a compressive non-linearity (see Fig. 7) that is characteristic of these curves. None of three lightness dependency functions, shown in Figure 7, is a good fit to these curves.

The effect of texture is to increase the tolerance thresholds above that of the uniform fields. This may be due to a masking of the mean lightness differences due to the additional spatial frequency components. The full-texture stimuli produce a larger masking effect because of the greater magnitude of contrast above the mean level. The effect of masking is dependent on both the spatial frequency constituents of the texture and the magnitude of these components.[24] It is likely that different types of textures will influence tolerance thresholds differentially. Textures with components that are tuned closely to the spatial frequencies of high visual sensitivity may reduce thresholds compared to uniform stimuli. In order to predict the effect of texture on tolerance, different types of textures need to be examined.

These experiments examined only the effect of texture to lightness tolerance. In the standard equations (CIE94 and CMC), the effect of texture has been implemented by increasing the contribution of lightness to the overall color difference metric by increasing the *l:c* ratio.[7] The *l:c* ratio also has been proposed as a way to control for changes in the magnitude of tolerance judgments and as a way to adjust for the scaling of judgements of acceptability rather than perceptibility.[7]

CONCLUSIONS

This experiment supports the findings of previous investigators that there is a lightness dependency in suprathreshold lightness tolerance.[6,9,20,22,25] In general, tolerances increase with increasing L* value. Texture also influences lightness tolerances. By simulating a texture of wound thread on a CRT it was shown that texture increases the tolerance thresholds. The data, plotted in CIELAB values, indicates that the form of the lightness dependency may be different in uniform and textured patterns.

When the data are plotted as a TVI plot, there is more regularity which may indicate that modeling lightness differences on the order of 1 L* unit may be better achieved if lightness is scaled differently. Small differences may be better represented in a logarithmically scaled space while perceptual scales, such as those used in uniform color spaces and color appearance spaces, may be better represented in spaces based on power functions.

The data show an influence of crispening. This effect of the background on the tolerance measurements has been said to have little influence[21,25] but has been noticed in the RIT-Dupont data set.[26] This size of this effect may depend on the experimental methods used in the particular experiment. More experimentation using different backgrounds is needed in order to measure the parametric influence of crispening on color difference thresholds.

In any event, crispening should be taken into account in the development of industrial tolerance equations in the future. Based on this experiment we can conclude that the absence of a lightness dependency term in CIE94 is a deficiency which should be addressed in future work towards a CIE standard. The CIE recognized, when putting CIE94 forward as an interim recommendation, that these effects, among many others, need more study before a standard is defined.[5]

This experiment explored only the effect of texture on lightness judgments. The effect of texture on hue and chroma and possible interactions is the focus of current research. The textures simulated in this experiment were neutral with variation only in lightness. Experiments are underway to see how this lightness variation in texture will affect tolerances in other directions in color space. More research is needed to investigate how textures with variation in the hue and chroma components influence tolerance thresholds.

The use of the CRT has allowed us to simulate textures. More sophisticated modeling of texture will allow better texture simulation on a CRT. These methods will allow for the rapid investigation of the effects of texture on industrial color tolerance. This research will lead to better equations for use in industrial color tolerancing. Also, as an extension to Berns,[7] it may be possible to optimize existing tolerance equations for particular commercial use based on pass-fail visual decisions and colorimetric data using texture simulation.

ACKNOWLEDGMENTS

This research was supported by the Munsell Color Science Laboratory Industrial Color Difference Consortium. The current members are 3M, Datacolor International, Bayer Corporation, Inter-Society Color Council, Macbeth, Society of Plastics Engineers, PPG, and Xerox Corporation.

1. Alman DH, Berns RS, Snyder GD, Larson WA. Performance testing of color-difference metrics using a color tolerance dataset. Color Res Appl 1989;14:139-151.
2. Berns RS, Alman DH, Reniff L, Snyder GD, Balonen-Rosen MR. Visual determination of suprathreshold color-difference tolerances using probit analysis. Color Res Appl 1991;16:297-316.
3. Qiao Y, Berns RS, Reniff L, Montag E. Visual determination of hue suprathreshold color-difference tolerances. Color Res Appl 1998;23:302:313.
4. Montag ED, and Berns RS. Visual determination of hue suprathreshold color-difference tolerances using CRT-generated stimuli. Color Res Appl 1999; 24:164-176.
5. CIE. Industrial colour-difference evaluation. Vienna: CIE Technical Report 116; 1995.
6. Clarke FJJ, McDonald R, Rigg B. Modification to the JPC79 colour-difference formula. J Soc Dyers and Colourists 1984;100:117-148.
7. Berns RS. Deriving instrumental tolerances from pass-fail and colorimetric data. Color Res Appl 1996;21:459-472.
8. Berns RS. The mathematical development of CIE TC 1-29 proposed color difference equation: CIELCH. AIC COLOUR 93, Proc. 7th Cong AIC. Budapest. 1993; C-19-1 –4.
9. McDonald R. Industrial pass/fail colour matching. Part III-Development of a pass/fail formula for use with instrumental measurement of colour difference. J Soc Dyers and Colourists 1980;96:486-497.
10. The MathWorks, Inc., Using MATLAB . Natick, MA: The MathWorks, Inc.; 1996.
11. Brainard DH. The Psychophysics Toolbox. Spatial Vision. 1997;10:443-446.
12. Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies, Spatial Vision. 1997;10:437-442.

13. Berns RS, Motta RJ, and Gorzynski ME. CRT colorimetry. Part I: Theory and practice. Color Res Appl 1993;18:299-314.

14. Berns RS. Methods for characterizing CRT displays. Displays 1996;16:173-182.

15. CIE. The relationship between digital and colorimetric data for computer-controlled CRT displays. Vienna: CIE Technical Report 122; 1996.

16. Heeger DJ and Bergen JR. Pyramid-based texture analysis/synthesis. Proc ACM Siggraph, August, 1995.

17. Finney DJ, Probit Analysis, 3rd Ed. Cambridge: Cambridge Univ. Press; 1971.

18. SAS Institute Inc. SAS User's Guide: Statistics. Cary, NC: SAS Institute, Inc.

19. Fairchild MD, Color Appearance Models, Reading Massachusetts: Addison-Wesley; 1998.

20. Kim D-H and Nobbs JH. New weighting functions for the weighted CIELAB colour difference formulae. AIC COLOUR 97. Proc 8th CongAIC. Kyoto. 1997;446-449.

21. Luo MR and Rigg B. BFD(l:c) colour-difference formula Part 1 – Development of the formula. J Soc Dyers and Colourists. 1987;103:86-94.

22. Luo R and Rigg B. Chromaticity-discrimination ellipses for surface colors. Color Res Appl. 1986; 11: 25-42.

23. Stiles WS. Increment thresholds and the mechanisms of colour vision. Documenta Ophthalmologica. 1949; 3: 138-163.

24. Legge GE and Foley JM. Contrast masking in human vision. J Opt Soc Am A. 1980; 70: 1458-1471.

25. Coates E, Fong KY and Rigg B. Uniform lightness scales. J Soc Dyers and Colourists. 1981;97:179-183.

26. Berns RS. Personal communication.

Figure 1. The four images used as templates for creating the sample images used in the experiment.

Figure 2. Histograms of the pixel values for the four template images. (a) L* histograms with a bin width of 1 L* unit. (b) C^*_{ab} histograms with a bin width of 1 C^*_{ab} unit. (c) h_{ab} histograms with a bin width of 1°.

Figure 3. Examples of the images used in the experiment. The top row shows full-texture samples with mean L* = 9, 39, 89 and reading from left to right. The middle row shows the half-texture samples of the same mean L* as the row above. The bottom row shows the uniform samples with the corresponding L* values.

Figure 4. The stimulus display on the CRT.

Figure 5. T50 values and 95% fiducial limits for neutral color centers centered near L* = 10, 20, 40, 60, 80 and 90. The arrow indicates the L* value of the background.

Figure 6. Results of the probit analysis for the full-texture samples at L* = 80. (a) Results from the first 20 observers. (b) Results from the second set of 24 observers. (c) Combined results.

Figure 7. Lightness dependency functions from three color tolerance equations: CMC, LCD, and BFD.

Figure 8. The data from Figure 5 are plotted as a TVI plot with the logarithm of the threshold, log(ΔY), as a function of the logarithm of the mean luminance of the color

centers, log(Y). The luminance of the background is indicated by the arrow. The dotted lines are the best linear fits to the three data sets.

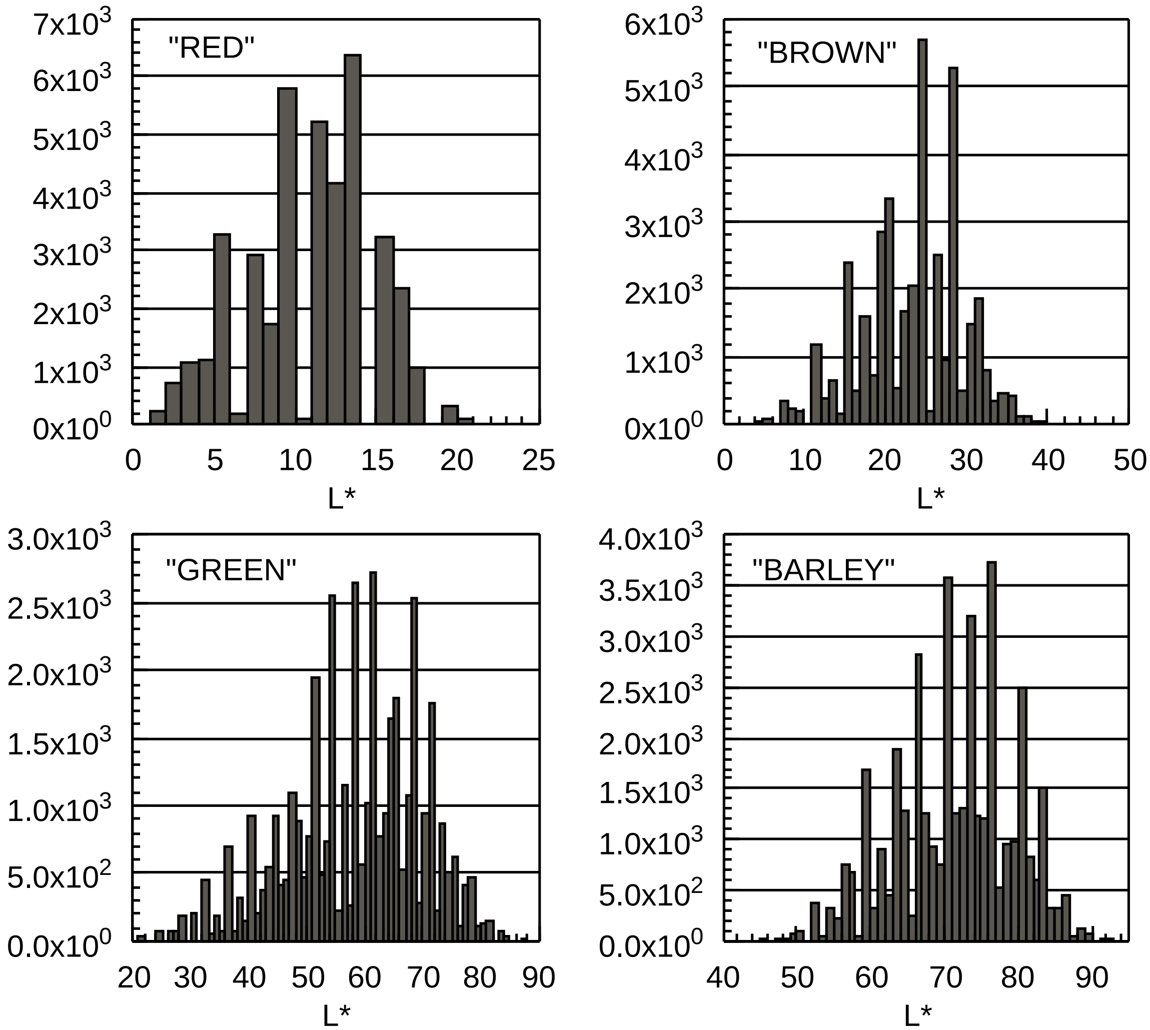
Figure 9. A family of curves based on the mean slopes of the fitted lines in Figure 7 are plotted as ΔL^* vs L^*. The original data and lightness dependency functions from Figure 5 are also plotted.

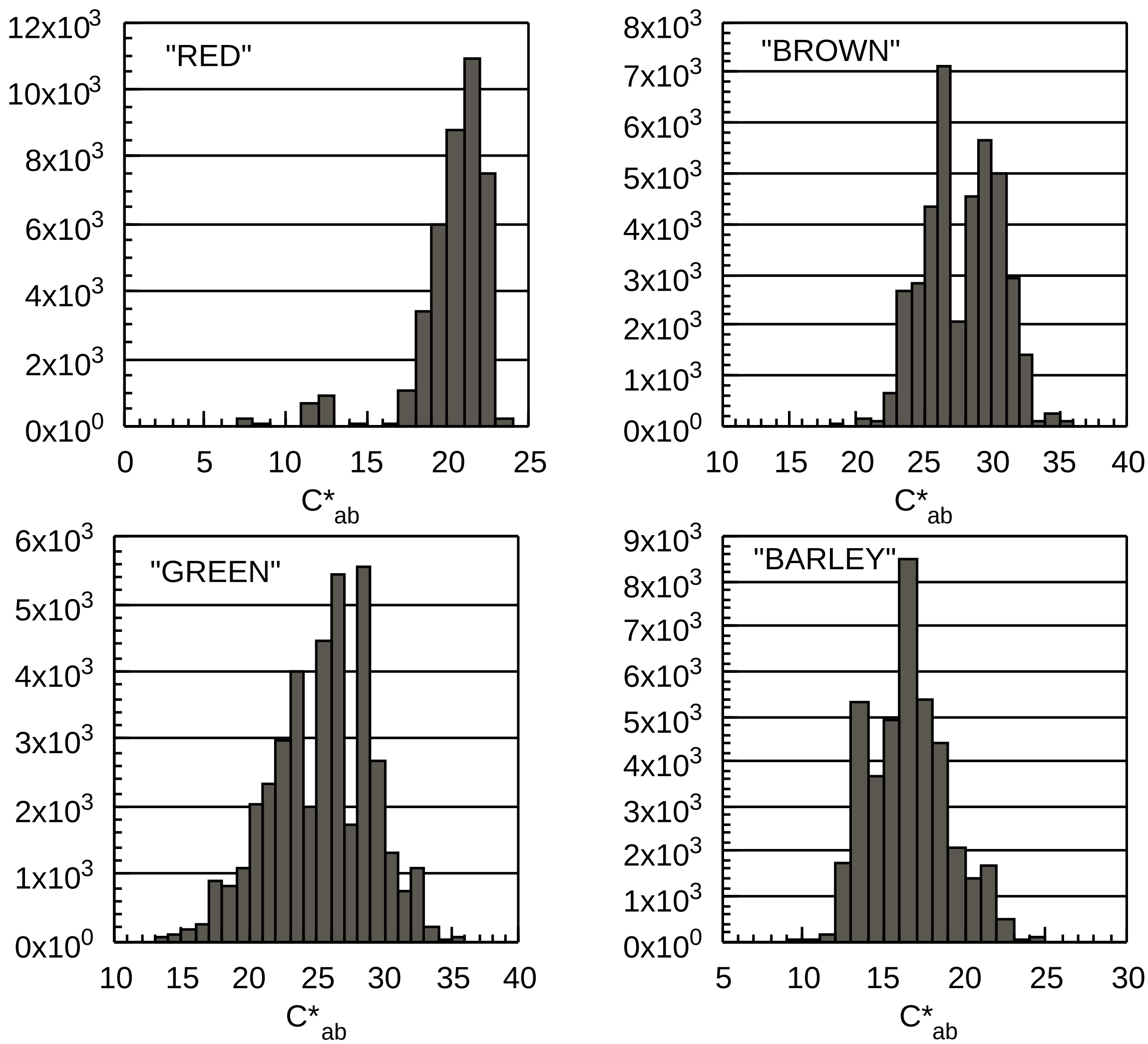
"Red"

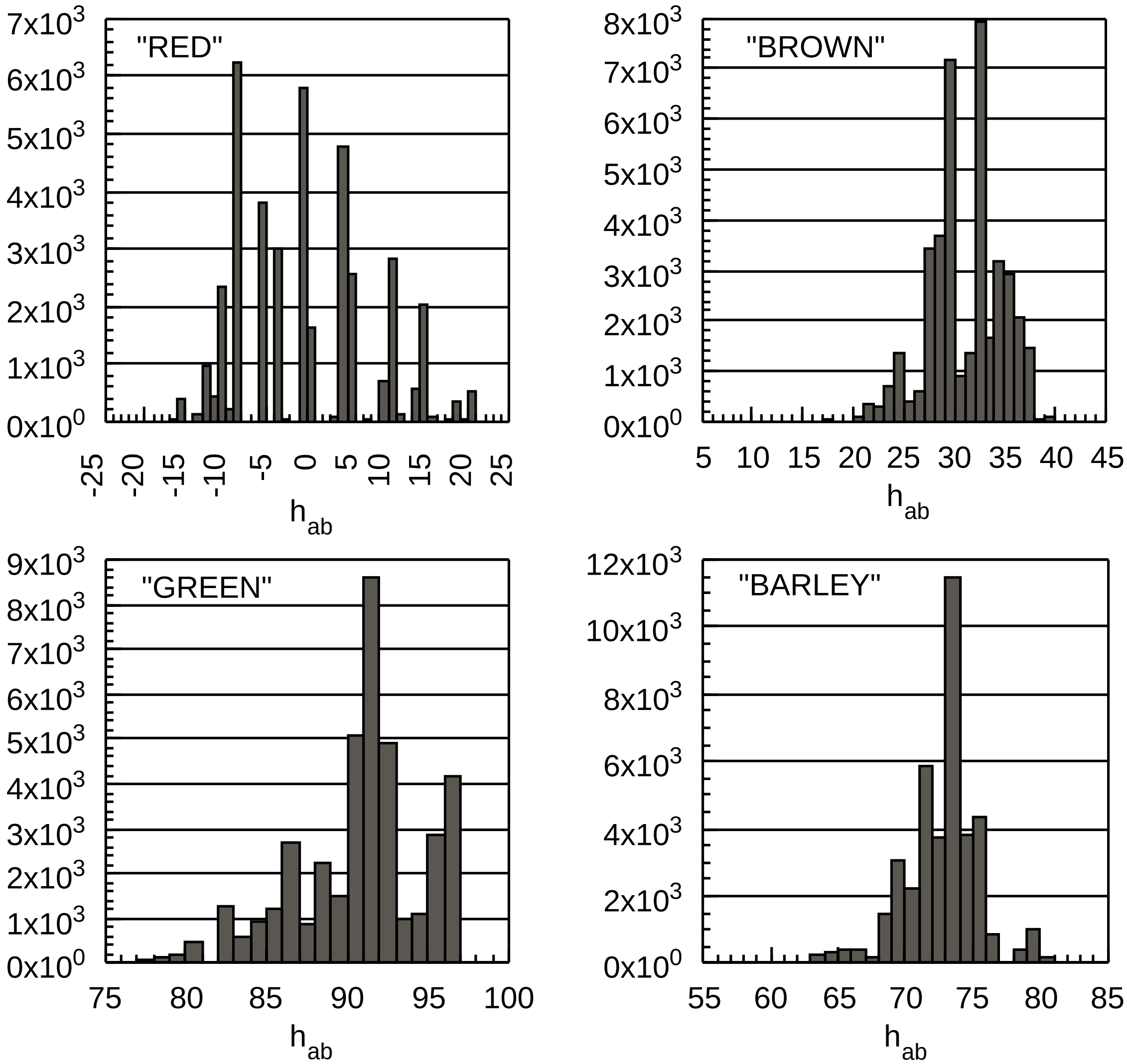
"Brown"

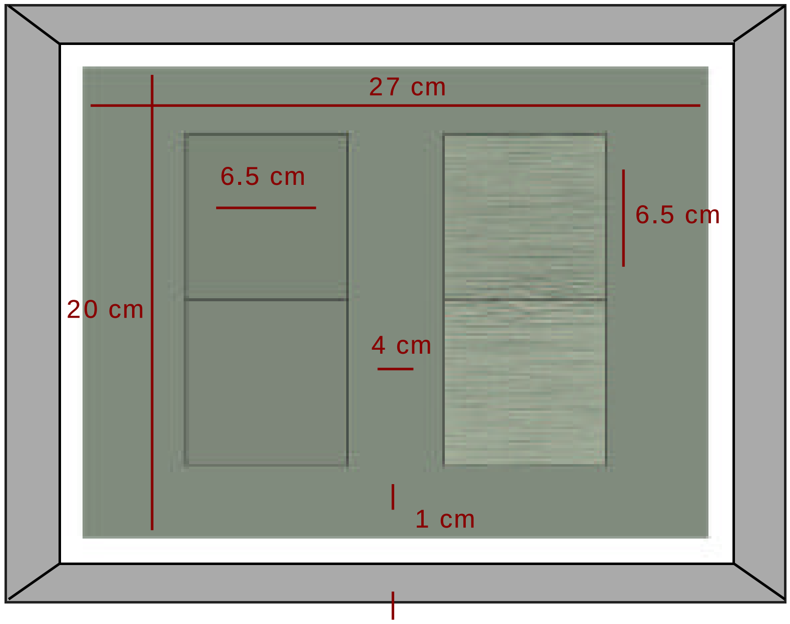
"Green"

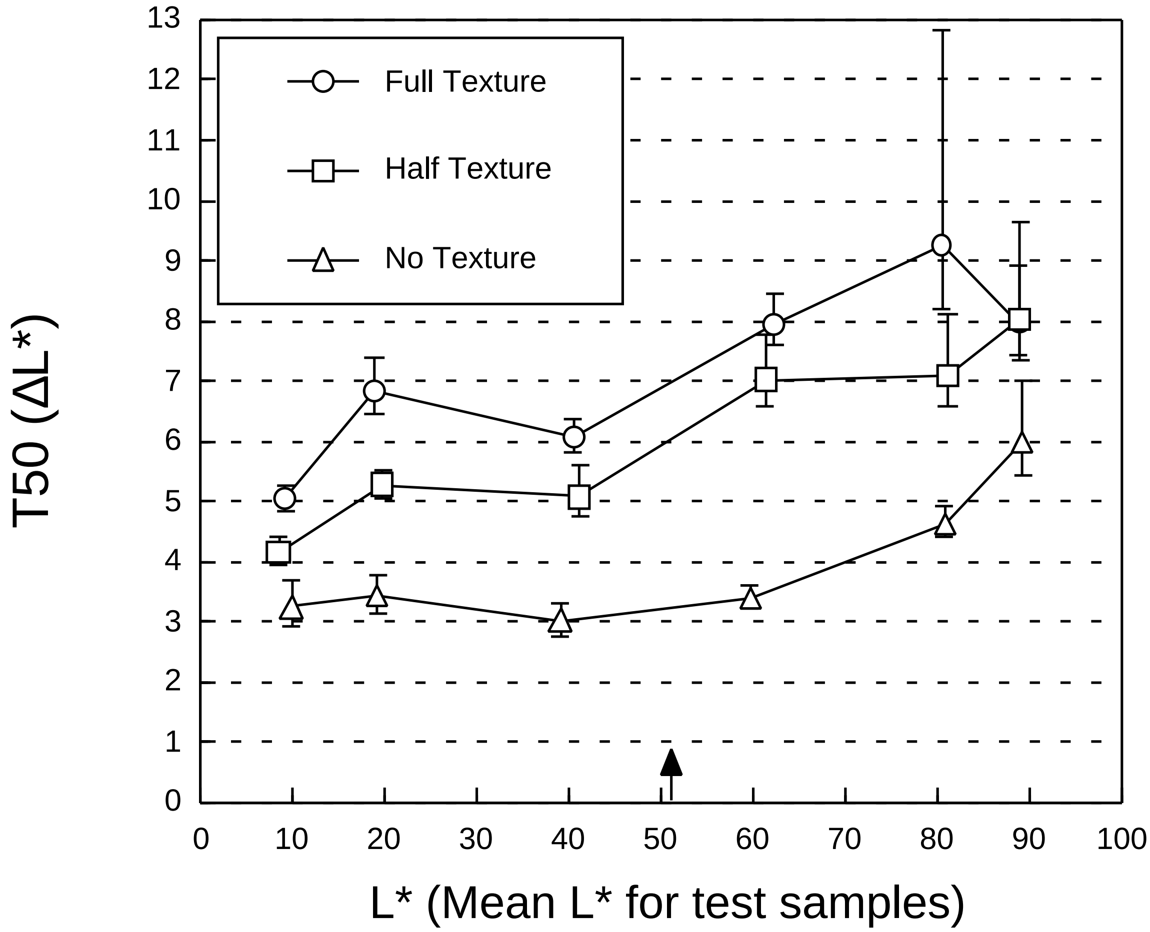
"Barley"


"RED"
"BROWN"
"GREEN"
"BARLEY"
L*
L*
L*
L*
7x10^3
6x10^3
5x10^3
4x10^3
3x10^3
2x10^3
1x10^3
0x10^0
0 5 10 15 20 25
6x10^3
5x10^3
4x10^3
3x10^3
2x10^3
1x10^3
0x10^0
0 10 20 30 40 50
3.0x10^3
2.5x10^3
2.0x10^3
1.5x10^3
1.0x10^3
5.0x10^2
0.0x10^0
20 30 40 50 60 70 80 90
4.0x10^3
3.5x10^3
3.0x10^3
2.5x10^3
2.0x10^3
1.5x10^3
1.0x10^3
5.0x10^2
0.0x10^0
40 50 60 70 80 90


"RED"
"BROWN"
"GREEN"
"BARLEY"
C*ab

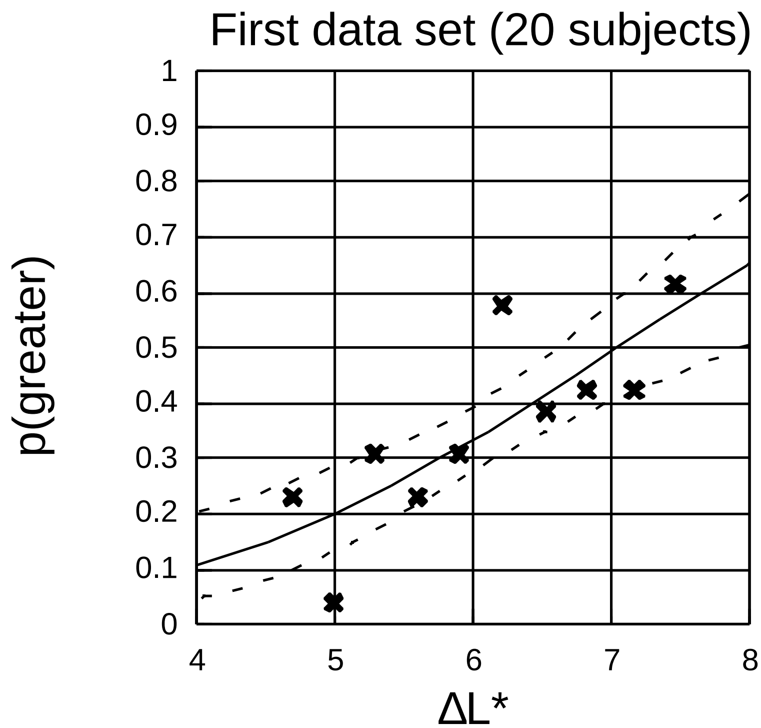

"RED"
"BROWN"
"GREEN"
"BARLEY"
h_ab
h_ab
h_ab
h_ab
0x10^0
1x10^3
2x10^3
3x10^3
4x10^3
5x10^3
6x10^3
7x10^3
8x10^3
9x10^3
10x10^3
12x10^3
-25 -20 -15 -10 -5 0 5 10 15 20 25
5 10 15 20 25 30 35 40 45
75 80 85 90 95 100
55 60 65 70 75 80 85


27 cm
6.5 cm
6.5 cm
20 cm
4 cm
1 cm



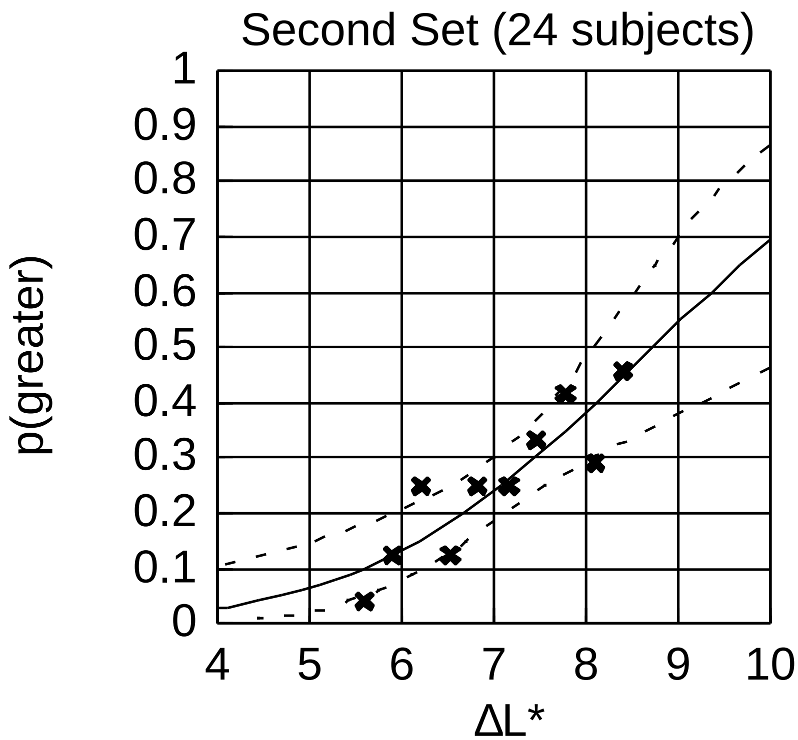
Full Texture
Half Texture
No Texture
T50 (ΔL*)
L* (Mean L* for test samples)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
0
10
20
30
40
50
60
70
80
90
100

6(a)



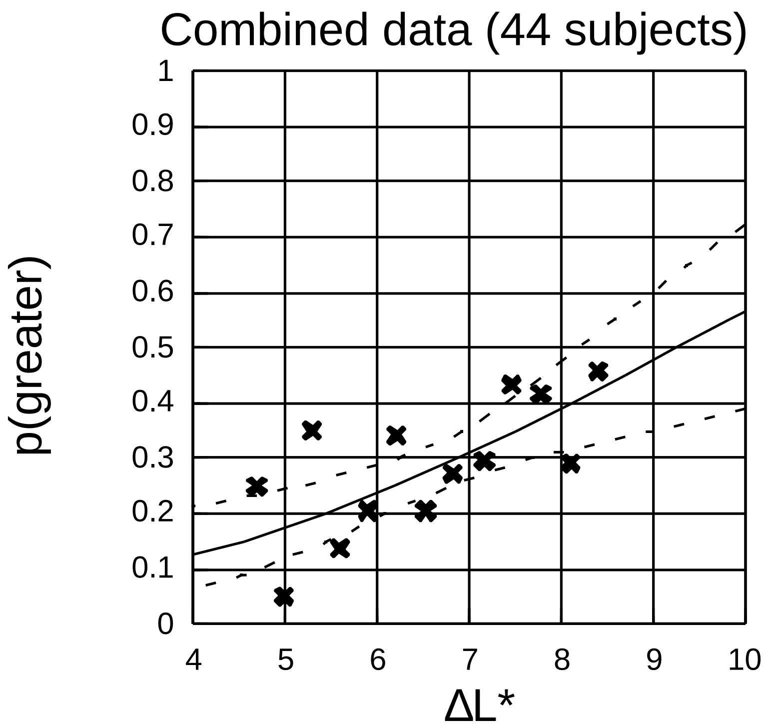
First data set (20 subjects)
p(greater)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
4
5
6
7
8
ΔL*

6(b)

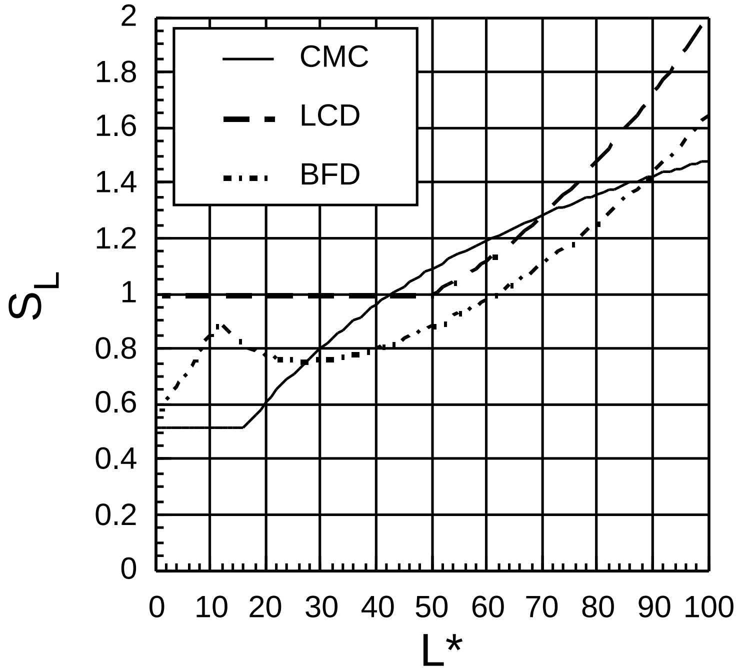


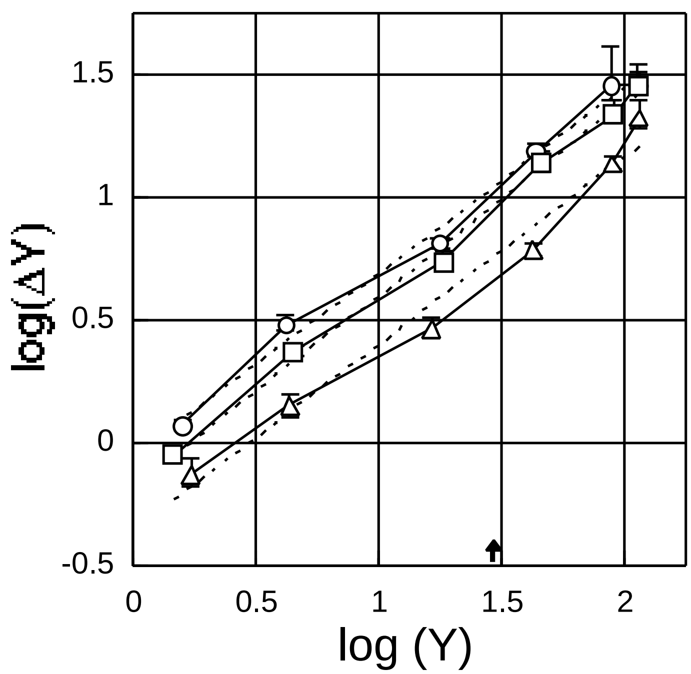
Second Set (24 subjects)
p(greater)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
4
5
6
7
8
9
10
ΔL*

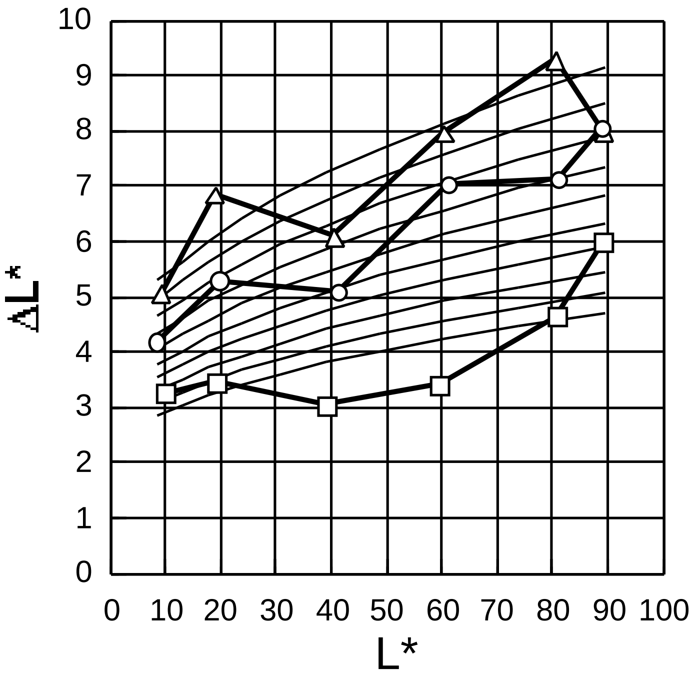
6(c)



Combined data (44 subjects)
p(greater)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
4
5
6
7
8
9
10
ΔL*


CMC
LCD
BFD
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
S_L
0 10 20 30 40 50 60 70 80 90 100
L*


1.5
1
0.5
0
-0.5
0
0.5
1
1.5
2
log(ΔY)
log (Y)



10
9
8
7
6
5
4
3
2
1
0
ΔL*
0 10 20 30 40 50 60 70 80 90 100
L*

Table I. Colorimetric values of the display.

White Border	$L^* = 100$, $x_{10} = 0.31$, $y_{10} = 0.33$, $Y = 138\ cd/m^2$		
Background	$L^* = 51.0$		
Anchor Pair	L^*	49.47	48.87
	a^*	0.58	0.10
	b^*	0.17	-0.53
	$\Delta E^*_{ab} = 1.03$, $\Delta L^* = 0.60$, $\Delta a^* = 0.48$, $\Delta b^* = 0.69$, $\Delta C^*_{ab} = 0.07$, $\Delta H^*_{ab} = 0.84$		

Table II. Values of the T50 thresholds for each lightness level and texture type. Fiducial limits and the range of ΔL* values of the combined set of test sample pairs are also included.

	Mean L*	T50 ΔL*	Lower Fiducial Limit	Upper Fiducial Limit	Range ΔL*
Full Texture	9.3	5.04	4.85	5.27	2.93 – 6.16
Half Texture	8.6	4.17	3.97	4.40	1.64 – 5.44
No Texture	10.0	3.25	2.92	3.68	0.92 – 4.58
Full Texture	19.0	6.84	6.47	7.39	3.20 - 7.79
Half Texture	19.9	5.28	5.05	5.54	1.85 – 6.87
No Texture	19.3	3.45	3.12	3.79	1.31 – 4.99
Full Texture	40.6	6.08	5.84	6.37	3.63 – 7.18
Half Texture	41.3	5.08	4.74	5.60	1.75 – 5.74
No Texture	39.2	3.03	2.76	3.31	1.22 – 4.52
Full Texture	62.4	7.94	7.60	8.44	4.39 – 8.70
Half Texture	61.4	7.02	6.58	7.76	3.07 – 7.27
No Texture	59.8	3.40	3.23	3.59	1.57 – 4.31
Full Texture	80.5	9.27	8.22	12.84	4.71 – 8.41
Half Texture	81.2	7.10	6.60	8.12	4.04 – 7.14
No Texture	80.9	4.63	4.40	4.94	2.53 – 5.53
Full Texture	88.9	7.96	7.44	8.94	4.30 – 7.92
Half Texture	89.0	8.03	7.33	9.63	3.96 – 7.94
No Texture	89.3	5.99	5.45	7.02	2.04 – 6.24