Title: Perceived image contrast and observer preference I. The effects of lightness, chroma, and sharpness manipulations on contrast perception

Authors: Anthony J. Calabria and Mark D. Fairchild

Author Info:

Anthony J. Calabria

Sun Chemical

631 Central Ave.

Carlstadt, NJ 07901

Mark D. Fairchild

Munsell Color Science Laboratory—Rochester Institute of Technology

54 Lomb Memorial Drive

Rochester, NY 14607

Abstract

A large-scale investigation into the perception of contrast in color images was performed. Psychophysical experiments were performed to determine the influence of image lightness, chroma, and sharpness transforms on perceived image contrast and observer preference. The influence of these transforms on perceived contrast was investigated separately by independent soft-copy, paired-comparison tests of contrast perception and image preference. The perception of contrast across transformations was also investigated, as was the perception of image contrast relative to the most preferred image manipulation. In all, four experiments of contrast perception and image preference were performed by at least thirty-two observers each.

Results of the *lightness*, *chroma*, and *sharpness-contrast experiments* indicate perceived image contrast is a function of multiple image characteristics as opposed to simply being a function of the dynamic range of image intensity. In the *lightness-contrast experiments*, images of identical white and black points were scaled to have significant differences in contrast based on their manipulations from the

original image. In the *chroma-contrast experiments*, images of identical lightness channels were scaled to have significant differences in perceived contrast due to relative chroma amount. In the *sharpness-contrast experiments*, images of identical white and black points were scaled to have significantly different levels of perceived contrast due to sharpness. In the *scale-linking experiment*, it was found that images of the above manipulations could be scaled similarly for perceived contrast. All scales of perceived contrast and image preference were found to be image independent among pictorial images, regardless of observer experience.

Introduction

Contrast is an image characteristic that is described as both a physical and perceptual attribute. In vision science, *contrast* defines the perception of spatial variation. The contrast of uniform patches on a uniform background, contrast of grayscale sinusoids, and contrast of colored text on a uniform colored background have been researched extensively in vision science. Contrast metrics dealing with these simple situations are typically a weighted ratio of measurable foreground and background characteristics (luminance, CIELAB L*, etc.). Although contrast defined as some ratio of luminances (such as Michelson contrast) may be appropriate when dealing with sinusoids or uniform patches, a luminance ratio may not necessarily correspond with perceived contrast in color images. This type of research is typically the base of contrast-sensitivity function (CSF) parameters utilized in image difference/quality models.[1]

When dealing with complex images, the use of the maximum and minimum luminance pixels may not coincide with perception of contrast over the entire image. Depending on the image subject matter, artifacts such as a speckle highlight could cause a luminance-based contrast metric to fall apart. A preferred version of an image may have the same luminance-based contrast parameters (min, max luminance pixels) as an undesirable overexposed or underexposed reproduction of that image. The term *contrast* in color imaging is commonly used as an overall image attribute. For the purposes of this research, image contrast is defined below.

Image contrast: the rate of change of the relative luminance of image elements of a reproduction as a function of the relative luminance of the same image elements of the original image.[2]

Image contrast is commonly defined in terms of an image tone reproduction curve (TRC). In image capture, the TRC represents the transformation from the actual scene luminance to the luminance of

the captured image. In image reproduction, the TRC often represents the luminance transform from an original image to its reproduction. Contrast is commonly thought of as the slope of the straight-line portion of the TRC between an image and its reproduction. The transformation in Figure 1 represents such a TRC. The term gamma (γ) is often used to describe the slope of this portion of a TRC on log-log coordinates.[3] This straight-line portion of a TRC represents the midtone region of the image, where there is a consistent separation of tone. By this definition image contrast is assumed to be independent of image chroma,

In the imaging community, research on perceived image contrast has been largely based on environmental aspects of an image or viewing system. Environmental aspects such as luminance level (Stevens Effect, Hunt Effect), surround (Bartleson-Breneman Equations) and degree of adaptation have been proven significant in image contrast perception.[4] To differentiate image contrast from perceived image contrast, the following definition is used in this research:

Perceived image contrast: the perception of the rate of change of the relative luminance of image elements of a reproduction as a function of the relative luminance of the same image elements of the preferred/ideal version of the image.

An image attribute that has gone largely ignored in image contrast studies is image chroma. It has been shown that brightness increases as a function of chroma (Helmholtz-Kohlrausch effect). In imaging, the Helmholtz-Kohlrausch effect would indicate a colorful image has perceptual brightness differences from its achromatic version. It is of interest whether this difference has an effect on the perception of image contrast. Also of interest is the effect of image sharpness on perceived image contrast.

The goal of this research is to contribute information relating to the perception of contrast in digital color images. This task is approached through the generation of a large-scale psychophysical data set of perceived contrast in color images. The effect of common achromatic manipulations on the perception of image contrast is investigated experimentally. Also investigated are the role of chroma and sharpness in image contrast perception. The ultimate goal of this research is the development of a metric of perceived image contrast for incorporation into both color image difference and color image quality models.

Experimental

For this research, soft-copy paired-comparison experiments were used in an attempt to scale the perception of image contrast for each of the test images and then link separate scales together. Experiments used to generate scales of preference and contrast perception are referred to as *lightness-contrast experiments*, *chroma-contrast experiments*, and *sharpness-contrast experiments*. Experiments used to link the results of the previous experiments referred to as the *scale-linking experiments*. Each set of experiments consisted of an *image preference test* and an *image-contrast perception test*.

Data Collection. A paired comparison graphical user interface (GUI) written in IDL 5.3 on a Macintosh platform was used to display images and record observations.[5] A two-mouse selection GUI was used, where clicks of the left or right mouse corresponded to a selection of their associated images. Images were loaded into computer memory and randomly generated pairs were displayed. Six frequency matrices were generated for each observer, for each test, corresponding to the six test images.

Device Characterization. Accurate display colorimetry was required for colorimetric image manipulations. An Apple Cinema Display monitor powered by a G4 Power Macintosh running OS 9 was characterized.[6] All measurements were taken in a darkened room commonly used for psychophysical experimentation. Luminance measurements were made using a Photoresearch Spectrascan 650 (PR 650) spectroradiometer. The PR 650 was given a one-hour warmup time. Colorimetric measurements were made using an LMT C1210 color meter. The colorimeter and monitor were given a two-hour warmup time. The monitor was set at its highest brightness level. Measurements were taken off a centered square patch (500 x 500 pixels) generated in IDL 5.3. The remainder of the display was filled with a medium gray background of RGB digital counts (128, 128, 128). Measurements of CIE XYZ tristimulus values were taken at 52 digital count levels (17 for each of the RGB channels individually, approximately spaced by a power of 2) three times and averaged.

The monitor characterization incorporated a 3x3 RGB-to-XYZ transformation of primaries matrix, a 3x1 flare matrix, and three one-dimensional lookup tables (LUTs). The complete monitor model had an accuracy of $\overline{\Delta E_{94}^{*}} = 0.38$ on an independent verification test of 27 distributed colors (combinations of digital counts 20, 80, 200).

Viewing Conditions. All experiments were performed in the same darkened room used for characterization measurements. Observers were seated approximately 24-36 inches from the monitor screen. This distance corresponds to approximately 45 pixels per degree (22.5 cpd). Observers positioned themselves comfortably at that distance, centered to the monitor screen.

Test Images. Six test images were selected from the Corbis® collection available at http://www.corbis.com. With respect to subject matter, five images were pictorial, and one was a medical image. Test images were named *wakeboarder*, *brainscan*, *pyramid*, *couple*, *veggies*, and *dinner*. Images with corresponding names and size in pixels are shown in Figure 2.

The initial colorimetry of these scenes was unknown; the monitor characterization was used to generate image colorimetry for manipulation and analysis. An assumption of this research is that image contrast can be perceived independent of the relationship between an image and its original scene. Therefore the relationship between the image data and their original scenes was not important. The unmanipulated RGB images were considered the originals but this description was not indicative of any particular image attributes such as preference or idealness.

Observers. A large group of observers, with a significant number of naïve observers was desired for these experiments. Observers consisted mainly of RIT faculty, staff, graduate and undergraduate students. Expert observers were considered to be observers with significant experience in the fields of color imaging and image perception. Naïve observers consisted of RIT undergraduate students, staff, and others unaffiliated with RIT. All observers had normal color vision. All observers performed both the *image-contrast perception* and *image-preference tests*. Observer information is shown in Table I.

Observer Instructions. In the *image preference test*, observers were shown pairs of images and given the following set of instructions:

Image Preference: You will be presented pairs of images. Your task is to select the image you prefer. If you prefer the image on the left, press the button on the left mouse. If you prefer the image on the right, press the button on the right mouse.

Observers were not instructed to on how to determine the image of preference, therefore were allowed to use whatever criterion they felt appropriate.

In the *image-contrast perception test*, observers were presented pairs of images and given the following set of instructions:

Image Contrast: You will be presented pairs of images. Your task is to select the image you perceive to be of higher contrast. If you perceive higher contrast in image on the left, press the button on the left mouse. If you perceive higher contrast in the image on the right, press the button on the right mouse.

Instructions were intentionally worded in such a way that observers familiar with image contrast were allowed to use their own definition. Observers who did not understand the concept of image contrast were given the following statement:

The image of higher contrast is the image you perceive to have more easily distinguishable objects.

Since observers are choosing the image of higher contrast in an image pair, it was only necessary to describe image contrast in a pair-wise manner. The difficulty of defining image contrast is avoided, and example images do not need to be presented.

Experiment I

Lightness-Contrast Experiments. Thirty-two observers participated in the *lightness-contrast experiments*, 16 were considered expert observers, and 16 were considered naïve. All observers completed both paired-comparison tests of *image preference* and *contrast perception*. Paired-comparison results were used to generate frequency matrices for the *image preference test* and *contrast perception test* (with respect to observer), for each of the six test images. Thurstone's law of comparative judgments, case V, with an incomplete data matrix was used for all scale generation.[7] This method uses a least-squares solution for unanimous observations. Interval scales of contrast and preference were generated for the twenty lightness-manipulated images, for each of the six test images.

Lightness-Contrast Experiments Images. Lightness channel transfer functions chosen were seven sigmoidal functions, four power functions, eight linear functions, and one histogram-equalization (see Figures 3-5). Of the seven sigmoidal functions, four were generated by cumulative normal functions of full-width at half-height 10, 15, 20, and 25 (images of these transfer functions are named *increase sigmoid 10, 15, 20, 25*). Three additional functions were generated by reflecting sigmoids *increase sigmoid 15, 20 and 25* about a line of slope 1 (images of these transfer functions are named *decrease sigmoid 15, 20, 25*).

Power functions applied to the image lightness channel were of magnitude 0.90, 0.95, 1.00, and 1.05 (named *pow 0.90, 0.95, 1.00, 1.05*). Linear functions applied to the image lightness channel were of slope 1.20, 1.15, 1.10, 1.05, 0.95, 0.90, 0.85, and 0.80 (named *lin 0.20, 0.15, 0.10, 0.05, -0.05, -0.10, -0.15, -0.20*). The histogram equalization method used was the *hist_equal* function in IDL 5.5. Lightness transfer functions were applied uniformly to the L* channels of the six test images, except for the histogram equalization performed on the *brainscan* image. Due to the significant black background of the *brianscan* image, the histogram equalization was limited to pixels of L* > 10.

Observer Expertise Comparison. A dual scaling[8] test was performed on observer results to determine any grouping due to observer experience. No such groupings were evident. Scales of image preference and perceived contrast were generated for the expert and naïve groups separately. Naïve observer scale values are shown as a function of expert observer scale values in Figures 6 and 7 for perceived contrast and image preference for the six test images. Linear fits to average scale values indicate the naïve observers had more variability in their observations than expert observers for both the *image-contrast perception test* (slope = 0.70, R^2=0.97) and *image preference test* (slope = 0.81, R^2=0.86). A slope of unity would indicate naïve observers had the same variation in their observations as expert observers.

Lightness-Contrast Experiments Analysis. In upcoming plots of scale values, error bars coinciding with 95% confidence limits of the mean were generated using Equation 1 unless otherwise noted**.**

$$95\%\ \text{confidence range} = S \pm \frac{1.38}{\sqrt{n}} \qquad (1)$$

In Equation 1, *S* represents the scale value (either preference or perceived contrast) and n represents the number of observers used in that test.

It is of interest whether the perception of image contrast is image independent, therefore scales of perceived contrast are shown with results from the six test images simultaneously. An important factor in the case of image independence for this research is image naturalness, where naturalness is defined as the conformity of the image to the ideas and expectations the observers have about the original scene at the time the picture was being taken.[9] Since the *brainscan* is a pseudocolored grayscale image, it has no naturalness. Keeping this in mind, the five pictorial images were analyzed as a separate group, as well as

with the complete set of six images. If the *brainscan* image behaves similarly to the pictorial images, perceived image contrast could be determined to be image and subject matter independent.

Image-Contrast Perception Test Results. Results of the *image-contrast perception test* are shown in Figure 8. Here, perceived contrast scales are plotted as a function of manipulation number (images with the same manipulation number had the same transforms applied). These results indicate differences in perceived image contrast can be scaled similarly since similar results were found for each image. Average scale values at each manipulation number are shown (scale values of the test image were averaged at each manipulation number). Perceived-contrast values fall within the 95% confidence limits from the mean for most image manipulation numbers. In Figure 8 manipulations of the *brainscan* image are shown to fall outside the confidence limits at higher levels of perceived contrast.

Image Preference Test Results. Despite having shown image independence among perceived contrast manipulations, image preference results (Figure 9) show the *brainscan* image falling outside confidence limits at several image manipulations. There is a higher level of image independence in Figure 10, where the analyses are limited to the pictorial images. However, preference scale values fall outside confidence limits of the mean at several image manipulation numbers.

Image Preference vs. Perceived Contrast Analysis. Plotting the scales of image preference against the scales of perceived-contrast of pictorial images (Figure 11) generates the psychophysical relationship between preference or quality and perceived image attribute.[9,10] This relationship, (also called the *preference-percept relationship*), is a non-monotonic, "inverted U" shape where image preference increases as a function of image percept, reaches a maximum, and then decreases. It appears the *veggies* and *pyramid* images have slightly different peaks from the other images. From Figure 10 it is shown that on average, there is no significant preference difference between the ten most preferred image manipulations. Therefore, differences in the peaks of Figure 11 are not significant. The preference-percept relationship is observed in Figure 12 by mean image preference plotted as a function of mean perceived image contrast. Although there is a range of perceived contrast values deemed equivalent for preference, the range of contrast values is wide enough for the "inverted U" trend to be apparent. This relationship was expected, and these results indicate perceived contrast may be image independent.

The preference-percept relationship is not present in the *brainscan* image results. These results have very little correlation between preference and contrast. Table II gives an indication as to the degree of difference between the medical image and pictorial images. The order in which images were rated for contrast and preference shows very little correlation. It is clear observers are judging some aspect of the medical image differently than the pictorial images.

Lightness-Contrast Experiments Conclusions. The results of the *lightness-contrast experiments* show that lightness manipulations of pictorial images resulting in differences in perceived contrast can be scaled in an image independent manner with a high level of image independence. Lightness-contrast-manipulated pictorial images can be scaled for preference with a lower level of image independence. It was also shown that a pseudocolored medical image was scaled similarly with pictorial images for perceived image contrast. There were significant differences between the pseudocolored image and the pictorial images when being scaled for preference. It is believed that since the medical image has no naturalness, observers used other criterion for judging preference than were used in the natural (pictorial) images. It is concluded from theses analyses that similar lightness-based image contrast manipulations of pictorial images can be perceived and scaled in an image independent manner.

Experiment II

Chroma-Contrast Experiments. In the *chroma-contrast experiments* 32 observers were used, 15 expert, 17 naïve. All observers completed both paired-comparison tests of *image preference* and *image-contrast perception*. Scale generation was performed using Thurstone's law of comparative judgments, case V, with an incomplete data matrix as described previously.

Chroma-Contrast Experiments Images. For the *chroma-contrast experiments*, the chroma (CIELAB C^*_{ab}) channel of the most preferred image from the *lightness-contrast experiments* (*decrease sigmoid 20*) was manipulated for each test image. The *decrease sigmoid 20* image digital counts were transformed to CIELAB L*C*h coordinates. Images were generated with 0%, 20%, 40%, 60%, 80%, 100% and 120% of original image chroma (images were named *0.0c, 0.20c, 0.40c, 0.60c, 0.80c, 1.00c,* and *1.20c*). The scaled chroma channels were then recombined with the lightness and hue channels of the original image. This procedure resulted in seven images of identical lightness and hue channels, with different chroma channels.

Observer Expertise Comparison. Scales of image preference and perceived contrast were generated for the expert and naïve groups singularly. Naïve observer scale values as a function of expert observer scale values are shown in Figures 13 and 14 for perceived contrast and image preference for the six test images. Linear fits to average scale values indicate the naïve observers had more variability in their observations than expert observers for both the *image-contrast perception test* (slope = 0.58, R^2=0.89) and *image preference test* (slope = 0.95, R^2=0.99).

Image Contrast Perception Test Results. Image independent scales of perceived contrast as a function of image chroma amount were generated (Figure 15). Although significant contrast differences due to chroma were shown to exist in images of identical lightness channel, the relationship between perceived contrast and relative chroma amount was not an increasing monotonic function as expected. The achromatic (0% chroma) image was perceived to have a significantly higher level of contrast than the image with 20% of the original image chroma. At the 20% chroma level and above, the perceived contrast increases monotonically with an s-shaped function. Further experimentation with a higher sampling of chroma at these lower levels would be needed to determine the nature of the relationship.

One possibility is that there is a discontinuity in the relationship between perceived contrast and chroma at the lowest chroma levels. Another possibility is that the difference in contrast perception of the achromatic image could be an empirically based environmental explanation of the interpretation of the scene.[11,12] Observers may perceive lower chroma images similarly to viewing a scene on a foggy day or when viewing through a screen. If objects appear to be less chromatic and distinguishable on a foggy day, lower contrast levels may be perceived.

Image Preference Test Results. As in the *contrast-perception test*, the achromatic images had higher preference scale values than the 20% chromatic images. Given the continued use of black-and-white photography (in advertising, newspapers, etc.) it was not surprising that observers would prefer an achromatic version of an image to a very low chroma version of that image. It is observed in Figure 16 the chroma-boosted (120%) *brainscan* image is relatively higher in preference than the pictorial images. The continued increase in preference could possibly be attributed to the unnatural subject matter. It is possible the decrease in preference of pictorial images at the 120% chroma level may be attributed to a decrease in

image naturalness. The *brainscan* image may have been more preferred with boosted chroma due to its unnatural subject matter.

Image Preference vs. Perceived Contrast Analysis. Plotting perceived contrast scale as a function of preference scale results in the preference-percept relationship observed previously. In Figure 17, mean image preference scale is shown as a function of mean perceived contrast scale for pictorial images. Preference increases as a function of perceived contrast to the 100% chroma level, then decreases. Observation of this trend is limited due to the lack of chroma-boosted images. The single outlier from the trend is the achromatic image. It is believed if more chroma-boosted images were included in the study, a more distinct relationship between preference and perceived contrast would be observed.

Chroma-Contrast Experiments Conclusions. Results of the *chroma-contrast experiments* have shown images of identical lightness channels can be scaled for perceived contrast as a function of relative chroma amount. It was also concluded that perceived contrast in achromatic images is judged differently than in chromatic images. The hypothesis of perceived contrast increasing monotonically with chroma was proven for chromatic images, however achromatic images did not fit this trend. In chromatic images, the relationship between chroma-contrast and preference follows a preference-percept relationship, where preference increases as a function of chroma-contrast to a point, then decreases.

The previous conclusions of the unnatural *brainscan* image being judged differently than the natural pictorial images were reinforced. Although a monotonic relationship was not observed, observer preference in chromatic versions of the *brainscan* image increased with chroma through the 120% chroma boosted image. Chroma-boosted versions of the pictorial images were all perceived to be of lower preference than the 100% chroma images, possibly due to a decrease in naturalness.

Experiment III.

Sharpness-Contrast Experiments. Thirty-two observers participated in the *sharpness-contrast experiments*, 15 were considered expert observers, and 17 were considered naïve. All observers completed both paired-comparison tests of *image preference* and *contrast perception*. Scale generation was performed using Thurstone's law of comparative judgments, case V, with an incomplete data matrix as described previously.

Sharpness-Contrast Experiments Images. For the *sharpness-contrast experiments*, the lightness channel of the most preferred image from the *lightness-contrast experiments* (*decrease sigmoid 20*) was manipulated for sharpness. The following procedure was used to in image sharpness manipulations. The *decrease sigmoid 20* image digital counts were transformed to CIELAB coordinates using the monitor forward model. Image lightness channels were removed from the color channels, passed through the inverse monitor model and written out as RGB images in TIFF format. The TIFF images were then opened in Adobe Photoshop®, where the *unsharp mask* filter was applied with radius = 2.0 and amount = 0, 25, 50, 75, 100, 150, 200, and 250. Sharpened images were saved in TIFF format and passed through the monitor forward model into CIELAB coordinates. The lightness channels of the sharpened images were then substituted for the unmanipulated lightness channels of the *decrease sigmoid 20* image. For each test image, this procedure resulted in eight image manipulations, each having a different lightness channel and identical chromatic channels (images were named *0sc, 25sc, 50sc, 75sc, 100sc, 150sc, 200sc* and *250sc*).

Observer Expertise Comparison. Scales of image preference and perceived contrast were generated for the expert and naïve groups separately. Naïve observer scale values are shown as a function of expert observer scale values in Figures 18 and 19 for perceived contrast and image preference scales. Linear fits to average scale values indicate the naïve observers had more variability in their observations than expert observers for the *image-contrast perception test* (slope = 0.83, R^2=0.99). In the *image preference test*, naïve observers had less variability than expert observers (slope = 1.06, R^2=0.94).

Image-Contrast Perception Test Results. Plots of perceived contrast vs. sharpness level were generated for all images (Figure 20). The relationship between perceived contrast and sharpness appears to be monotonic and image independent. Assumptions as to a linear relationship between contrast and sharpness cannot be determined based on these data. It is unknown if there is a linear relationship between the values used in the Photoshop *unsharp mask* filter and the nature of the filter.

Image Preference Test Results. Plots of image preference vs. sharpness level were generated (Figure 21 for pictorial images only). On average, preference was shown to increase with sharpness to a point, and then decrease. Manipulations of the *brainscan* image were judged differently than the pictorial images and therefore are not shown. The preference of the *brainscan* peaked at the 150sc level, as opposed to the 25sc level of the pictorial images. Preference as a function of sharpness produced the most image dependent

results encountered in this research. Image dependence was observed between pictorial images. A possible reason for this image dependence could be the application of the *unsharp-mask* filter. Frequency analysis of sharpened images indicates different regions of the power spectra were manipulated for each of the test images. It is of interest to determine if the image dependence aspect of preference vs. sharpness level would decrease if similar frequency bands were enhanced for each test image.

Perceived Image Contrast vs. Image Preference Analysis. Plots of image preference vs. perceived contrast were generated. Figure 22 illustrates sharpness-contrast manipulations follow the preference-percept relationship for the five pictorial test images. Unlike the preference-percept relationships observed in the *lightness-contrast* and *chroma-contrast experiments*, there is much more variability in results of the *sharpness-contrast experiments* due to variability exhibited in the image preference results.

Sharpness-Contrast Experiments Conclusions. Results of the *sharpness-contrast experiments* have shown sharpness manipulations of image lightness channels can be scaled for perceived contrast as a function of sharpness. It can therefore be concluded the perception of image contrast is a monotonic function of image sharpness, independent of image content. On average, the relationship between sharpness-contrast and image preference follows the preference-percept relationship, where preference increases as a function of sharpness-contrast to a point, then decreases. The previous conclusions of the unnatural *brainscan* image being judged differently than the natural pictorial images were again reinforced. Although contrast was perceived similarly in the *brainscan* image as the pictorial images, it was judged differently in the *image-preference experiments*. The most preferred *brainscan* manipulation was at a much higher sharpness-contrast level than the most preferred pictorial images.

Experiment IV.

Scale-Linking Experiments. The goal of the *scale-linking experiments* was to link the previously generated scales of perceived image contrast and image preference from the *lightness-contrast*, *chroma-contrast*, and *sharpness-contrast experiments*. This is done because the interval scales developed in the previous experiments are independent of one another and cannot be linked unless stimuli are included in the separate experiments. If possible, this would give an indication of what factors a metric of perceived image contrast should take into consideration.

Thirty-two observers participated in the *scale-linking experiments*, 16 were considered expert observers, and 16 were considered naïve. All observers completed both paired-comparison tests of *image preference* and *contrast perception*. Scale generation was performed using Thurstone's law of comparative judgments, case V, with an incomplete data matrix as described previously.

Scale-Linking Experiments Images. To link the three previously generated scales of perceived contrast and image preference, images from each scale were compared to each other in an independent experiment. No new images were generated for this experiment. Based on the results of the previous experiments, only the five pictorial images were used. Pairs were only compared between manipulations of like subject matter. Twelve images were chosen for the *scale-linking experiments*. The five images chosen from the *lightness-contrast experiments* were: increase sigmoid 10, increase sigmoid 25, decrease sigmoid 20, linear –0.150, and linear 0.150. Three images chosen from the *chroma-contrast experiments* were: 1.20c, 0.60c, and 0.20c. Four images chosen from the *sharpness-contrast experiments* were: 250sc, 150sc, 75sc, 25sc.

Observer Expertise Comparison. Scales of image preference and perceived contrast were generated for the expert and naïve groups singularly in Figures 23 and 24. Linear fits to average scale values indicate the naïve observers had more variability in their observations than expert observers for both the *image-contrast perception test* (slope = 0.63, R^2=0.96) and *image preference test* (slope = 0.73, R^2=0.93).

Image Contrast Perception Test Results. Scales of perceived image contrast are shown in Figure 25 for the five pictorial images. The linked contrast perception scales appear image independent. Image scale values fall within 95% confidence limits of the mean at nearly every contrast level.

Image Preference Test Results. Scales of image preference are shown in Figure 26 for the five pictorial images. The image dependency shown in Figure 26 was expected given the results of the previous *image preference tests*. Scales from the *dinner* image seem to be the only poor fit to the mean.

Perceived Image Contrast vs. Image Preference Analysis. Plots of mean image preference vs. mean perceived contrast (Figure 27) illustrate the *scale-linking experiment* perceived-contrast scale follows the preference-percept relationship. Figure 27 reveals manipulations of image lightness, chroma, and sharpness can be scaled for perceived contrast similarly. These results lead to the hypothesis that perceived contrast can be a described as a function of lightness, chroma, and sharpness manipulations.

Linking of Perceived Image Contrast and Image Preference Scales. Having demonstrated the ability to scale contrast manipulations of lightness, sharpness, and chroma scales of contrast-perception and image preference in an image independent manner, independently generated scales from the *lightness*, *sharpness*, and *chroma-contrast experiments* were linked. Using a least-squares (pseudoinverse) solution, coefficients of a linear transformation were generated for scales of perceived contrast and image preference from the *lightness*, *chroma*, and *sharpness-contrast experiments*. This is shown in matrix notation in Equations 2 and 3.

$$S_2 = S_1 b \tag{2}$$

$$b = \left(S_1^T S_1\right)^{-1} S_1^T S_2 \tag{3}$$

Engledrum[9] has shown the coefficients b_0 and b_1 of vector $\boldsymbol{b}$ are calculated by the pseudoinverse solution between scale values of samples on scale 1 (S_1) and the scale values of those same samples on scale 2 (S_2). In the case of this experiment, scales of preference and perceived contrast from the *lightness-contrast*, *chroma-contrast*, and *sharpness-contrast experiments* would be represented by the scale S_1, and the scales of preference and perceived contrast generated in *scale-linking experiments* would be represented by the scale S_2. Each scale of contrast and preference generated a unique set of coefficients for the transformation to the linked scale (Table III).

Equations 4 to 6 were used to transform *lightness*, *sharpness* and *chroma-contrast* scales to the scales of contrast perception and image preference generated in the *scale-linking experiments*.

$$\begin{bmatrix} C_{2L} \\ \vdots \\ C_{2L} \end{bmatrix} = \begin{bmatrix} 1 & C_{1L} \\ \vdots & \vdots \\ 1 & C_{1L} \end{bmatrix} \begin{bmatrix} b_{0L} \\ b_{1L} \end{bmatrix} \tag{4}$$

$$\begin{bmatrix} C_{2C} \\ \vdots \\ C_{2C} \end{bmatrix} = \begin{bmatrix} 1 & C_{1C} \\ \vdots & \vdots \\ 1 & C_{1C} \end{bmatrix} \begin{bmatrix} b_{0C} \\ b_{1C} \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} C_{2S} \\ \vdots \\ C_{2S} \end{bmatrix} = \begin{bmatrix} 1 & C_{1S} \\ \vdots & \vdots \\ 1 & C_{1S} \end{bmatrix} \begin{bmatrix} b_{0S} \\ b_{1S} \end{bmatrix} \tag{6}$$

Original arrays of contrast scale values C_1 (subscripted with L (lightness), C (chroma) and S (sharpness)) are transformed by the appropriate coefficients b_0 and b_1 shown above, to the contrast values of scale C_2 (subscripted similarly as C_1).

Since image independence of perceived-contrast has been established in previous experiments, the mean scale values of the lightness, chroma, and sharpness scales were linked, rather than the scales associated with each test image. Despite some image dependence with respect to preference, mean scale values of image preference were used as well. Mean scale values were transformed to the scales generated in the *scale-linking experiments*, along with the values of their associated error bar values. The actual scale values of the 95% limits were transformed as if they were scale values themselves. This was done to ensure significant differences present in the individual scales were maintained in the linked scale. Propagation of confidence limits to the linked scale resulted in 95% confidence limits of three different magnitudes.

It is observed from Figure 28 the mean scale values can be transformed to linked scale values while maintaining significant differences present in individual scales. Data labels in Figure 28 correspond to the order that manipulation was on the individual scale (CS01 had the highest contrast value on the sharpness-contrast scale, CC03 had the third highest value on the chroma-contrast scale).

The three highest perceived contrast manipulations were all sharpened, with the 250sc image judged as highest in contrast overall. The lowest contrast image was determined to be the 20% chroma image. Interestingly, perceived contrast of this image is determined to be not significantly different from the *gma_0.900* image. This is worth noting because the white point of the *gma_0.900* image is approximately L* = 63. The 20% chroma image has a white point of L* = 100. Since both images have black points of L* = 0, it is possible to refute claims that image contrast is solely a function of image black and white points. A linked image preference scale of mean scale values was generated similarly (Figure 29). Linking the scales resulted in the image 25sc as the most preferred image manipulation on average. The least preferred image manipulation, on average, was the 20% chroma image. Despite several images falling within the confidence limits of others, the scales cover a large enough area to allow for several levels of significant difference. Plotting linked image preference scales as a function of linked perceived contrast scales (Figure 30) resulted in a fairly complete preference-percept relationship curve. The

perceived contrast range covered by the linked scale is greater than those generated previously. Chroma-manipulated images are observed at the lower-contrast side of the curve, sharpened images are towards the higher-contrast side of the curve. The lightness-manipulated images cover a wider range of perceived contrast than the chroma or sharpness manipulated images.

Scale-Linking Experiments Conclusions. It has been proven that images having been manipulated for lightness-contrast, sharpness-contrast and chroma-contrast can be scaled similarly for perceived contrast. In addition, scales of perceived contrast can have been linked using a least-squares solution. Both the scaling and linking have been done in an image independent manner. Similar scales can be generated for image preference. Relating scales of preference as a function of perceived contrast results in the standard preference-percept relationship confirming the idea that perceived contrast in an actual perceptual quality of images that can be scaled or quantified.

Conclusions:

Perceived image contrast was determined to be scalable with respect to lightness, chroma, and sharpness manipulations. Perceived image contrast scales were image independent between five pictorial images. Significant contrast differences between images of identical white and black points were perceived, demonstrating that image white and black points do not solely determine image contrast. Significant image contrast differences were found between full color images and their achromatic versions, thus demonstrating that perceived image contrast is a function of image chroma information. It was also shown that the perceived contrast of achromatic images is higher than perceived contrast of very low-chroma images.

It was learned that perceived contrast in pictorial (natural) images could be gauged independently of image content. Different pictorial images of similar image manipulations can be scaled for contrast similarly. It was also learned that pseudocolored (unnatural) images could be scaled for perceived contrast similarly to pictorial images. Image preference of pseudocolored images was judged differently than image preference of pictorial images. Naïve observers have been shown to perceive contrast similarly to expert observers. This information is of significance to the color imaging community and should be incorporated into further studies, as necessary.

1. Johnson, G.M., On contrast sensitivity in an image difference model, *Proceedings from IS&T's 2002 PICS Conference*, (2002).

2. Fairchild, M.D., Considering the Surround in Device-Independent Color Imaging, *Color Research and Application*, **20**(6), 352-363, (1995).

3. Stroebel, L., Compton, J., Current.I, Zakia, R., *Photographic Materials and Processes* Focal Press, Boston, (1986).

4. Farichild, M.D., *Color Appearance Models*, Addison-Wesley, Reading, MA, (1998).

5. Johnson, G.M. Sharpness Rules, Proceedings from the 8th IS&T/SID Color Imaging Conference, (2000).

6. Fairchild, M.D., Wyble, D.R., Colorimetric Characterization for the Apple Studio Display (Flat Panel LCD), *Munsell Color Science Laboratory Technical Report*, (1998).

7. Engeldrum, P.G., *Psychometric Scaling*, Imcotek Press, Winchester, MA, (2000).

8. Nishisato, S. & Nishisato, I., *Dual Scaling in a Nutshell*, MicroStats, (1994).

9. De Ridder, H., Naturalness and Image Quality: Saturation and lightness variation in color images, *Journal of Imaging Science and Technology*, **40**(6), 487-493, (1996).

10. Engeldrum, P.G., Extending Image Quality Models, *Proceedings from IS&T's 2002 PICS Conference*, 65-69, (2002).

11. Lotto, R. B., Purves, D., An empirical explanation of color contrast, *Proceedings of the National Academy of Sciences of the USA*, **97**(23), 12389-12932, (2000).

12. Purves, D., Lotto, R.B., Williams, S.M., Nundy, S., Yang, Z. Why we see things the way we do: evidence for a wholly empirical strategy of vision, *Phil. Trans. R. Soc. Lond.*, **356**, 285-297, (2001).

Figure 1. Example of a “sigmoidal” tone reproduction curve (TRC).

Figure 2. Test images selected for this research with their corresponding size in pixels.

Figures 3,4,5. Examples of linear, sigmoidal, and power lightness-transfer functions.

Figure 6. Lightness-contrast experiment expert observer perceived contrast scale vs. naïve observer perceived contrast scale.

Figure 7. Lightness-contrast experiment expert observer image preference scale vs. naïve observer image preference scale.

Figure 8. Perceived image contrast scale vs. manipulation number (See Table 4-1 for manipulation numbers).

Figure 9. Image preference scale vs. manipulation number (See Table 4-2 for manipulation numbers).

Figure 10. Image preference scale vs. manipulation number (pictorial only, see Table 4-2 for manipulation numbers).

Figure 11. Image preference scale vs. perceived lightness-contrast scale (pictorial only)

Figure 12. Mean preference scale vs. mean perceived contrast scale (pictorial only)

Figure 13. Chroma-contrast experiment expert observer perceived contrast scale vs. naïve observer perceived contrast scale.

Figure 14. Chroma-contrast experiment expert observer image preference scale vs. naïve observer image preference scale.

Figure 15. Perceived contrast scale vs. chroma scalar for all six test images.

Figure 16. Image preference scale vs. chroma scalar for the six test images.

Figure 17. Mean preference scale vs. perceived contrast scale (pictorial only).

Figure 18. Sharpness-contrast naive observer perceived contrast scale vs. expert observer perceived contrast scale.

Figure 19. Sharpness-contrast naive observer image preference scale vs. expert observer image preference scale.

Figure 20. Perceived contrast scale vs. sharpness level for the six test images.

Figure 21. Image preference scale vs. sharpness level for the five pictorial images only (see Table 6-2 for sharpness level).

Figure 22. Mean preference scale vs. sharpness-contrast scale for the five pictorial images only.

Figure 23. Scale-linking experiment expert observer perceived contrast scale vs. naïve observer perceived contrast scale.

Figure 24. Scale-linking experiment expert observer image preference scale vs. naïve observer image preference scale.

Figure 25. Linking experiment perceived image contrast scale (See Table 4-1 for image numbers).

Figure 26. Linking experiment image preference scale (See Table 4-2 for image numbers).

Figure 27. Linking experiment mean perceived contrast scale vs. mean preference scale.

Figure 28. Linked perceived-contrast scale in order from highest perceived contrast to lowest perceived contrast with 95% confidence limits.

Figure 29. Linked image preference scale in order from most preferred manipulation to least preferred with 95% confidence limits.

Figure 30. Linked mean preference scale vs. linked mean perceived contrast scale.

Table I. Observer statistics from the five experiments. For ethnicity, C=Caucasian, A=Asian, ME=Middle Eastern, H=Hispanic.

Table II. Order of perceived contrast and image preference for the pictorial images and the medical image (1 = highest scale value, 20 = lowest scale value).

Table III. Linear transform parameters used for linking the lightness, chroma, and sharpness-contrast scales.

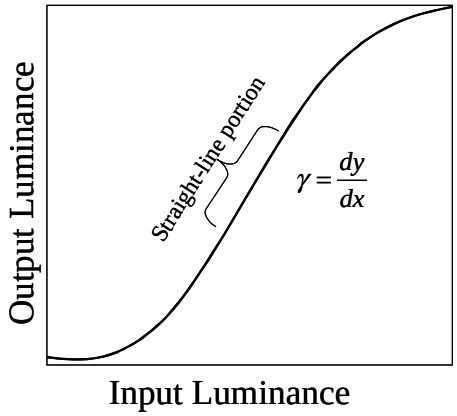
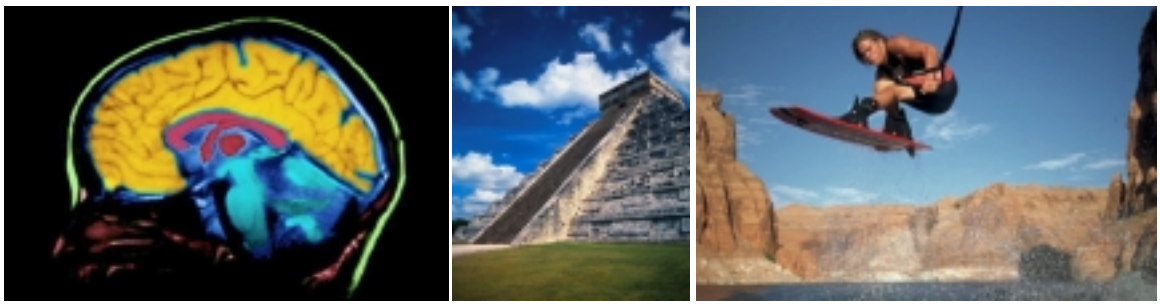


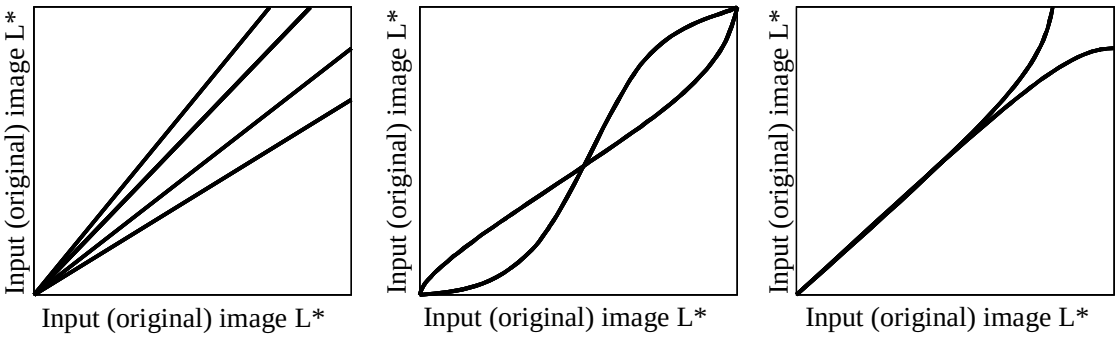
Figure 1. Example of a "sigmoidal" tone reproduction curve (TRC).



brainscan (640 x 424) pyramid (384 x 480) wakeboarder (640 x 410)

couple (320 x 480) fruits & veggies (614 x 480) dinner (384 x 480)

Figures 2. Test images selected for this research with their corresponding size in pixels.



Figures 3,4,5. Examples of *linear*, *sigmoidal*, and *power* lightness-transfer functions.

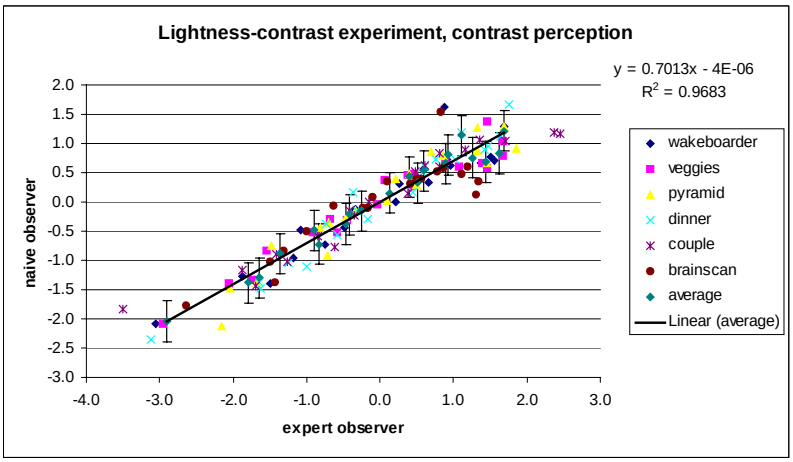


Figure 6. Lightness-contrast experiment expert observer perceived contrast scale vs. naïve observer perceived contrast scale.

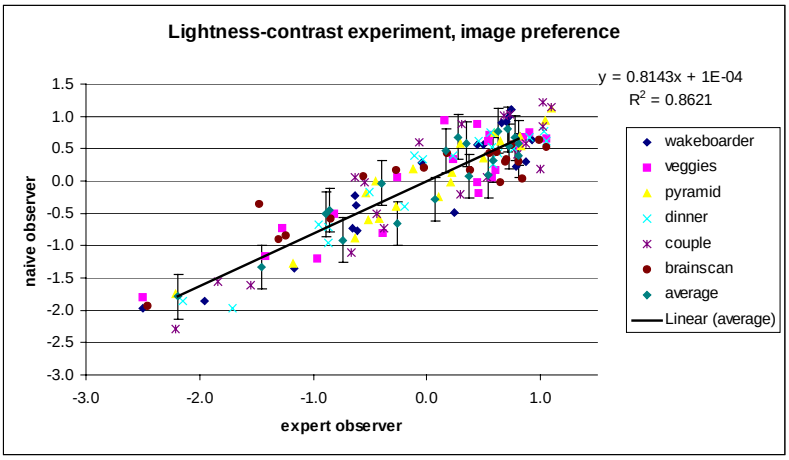


Figure 7. Lightness-contrast experiment expert observer image preference scale vs. naïve observer image preference scale.

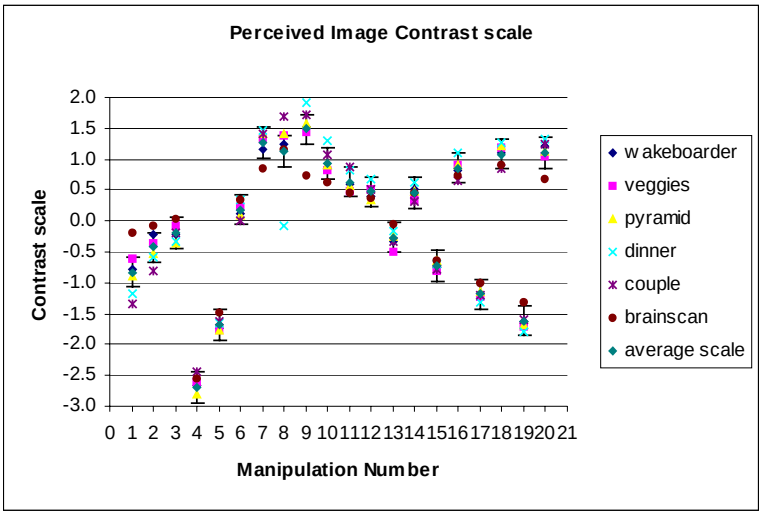


Figure 8. Perceived image contrast scale vs. manipulation number (See Table 4-1 for manipulation numbers).

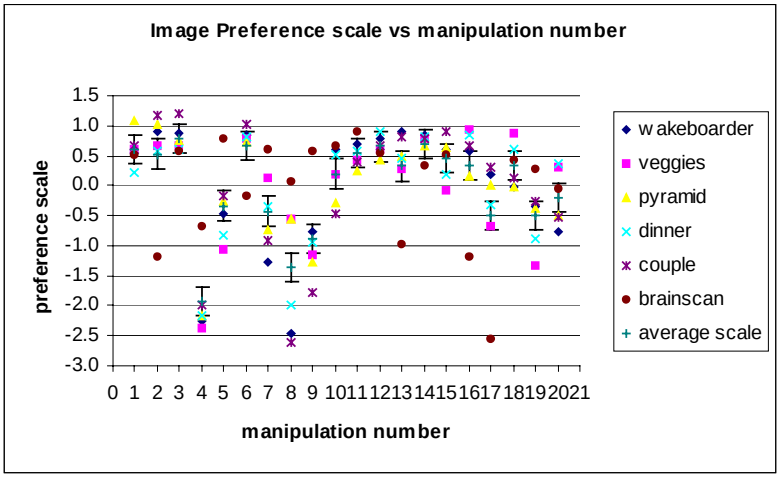


Figure 9. Image preference scale vs. manipulation number (See Table 4-2 for manipulation numbers).

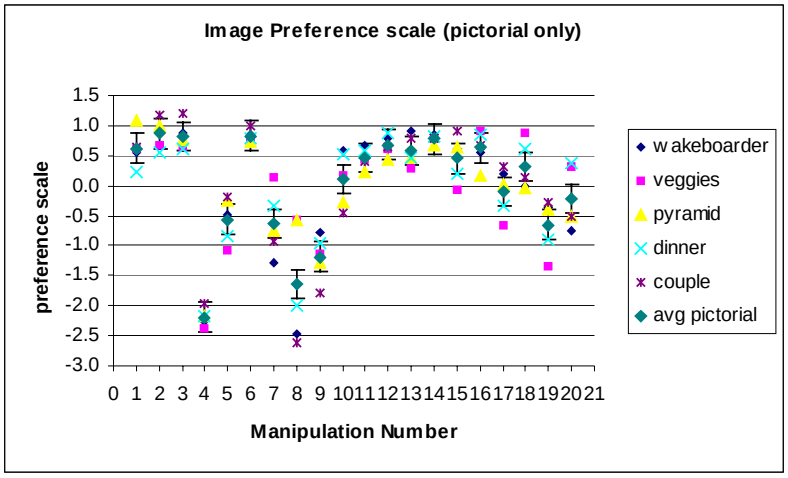


Figure 10. Image preference scale vs. manipulation number (pictorial only, see Table 4-2 for manipulation numbers).

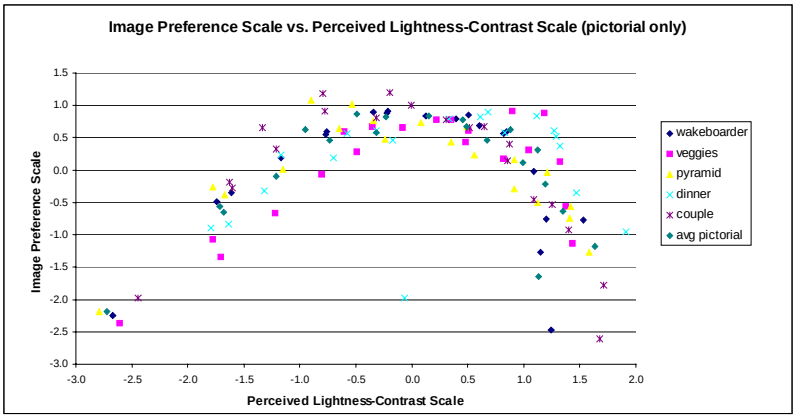


Figure 11. Image preference scale vs. perceived lightness-contrast scale (pictorial only)

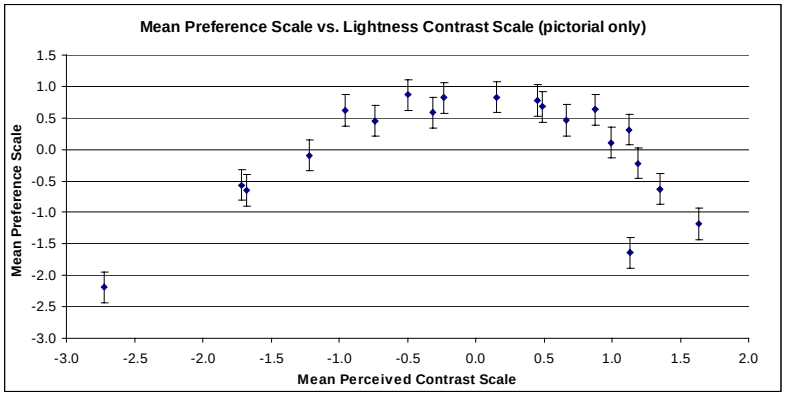


Figure 12. Mean preference scale vs. mean perceived contrast scale (pictorial only)

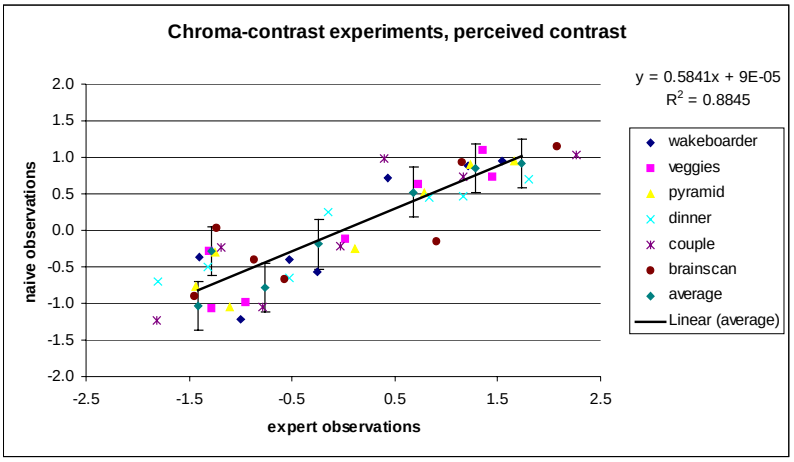


Figure 13. Chroma-contrast experiment expert observer perceived contrast scale vs. naïve observer perceived contrast scale.

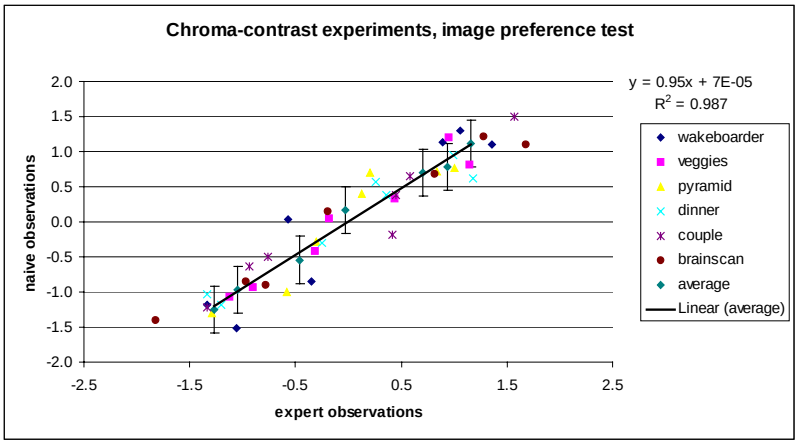


Figure 14. Chroma-contrast experiment expert observer image preference scale vs. naïve observer image preference scale.

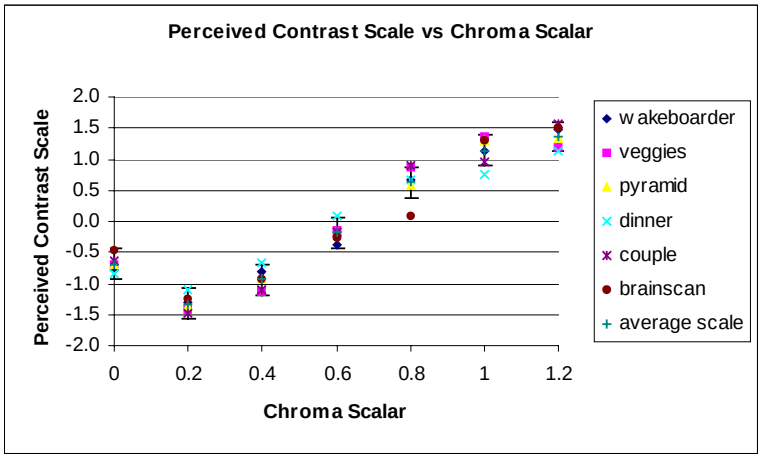


Figure 15. Perceived contrast scale vs. chroma scalar for all six test images.

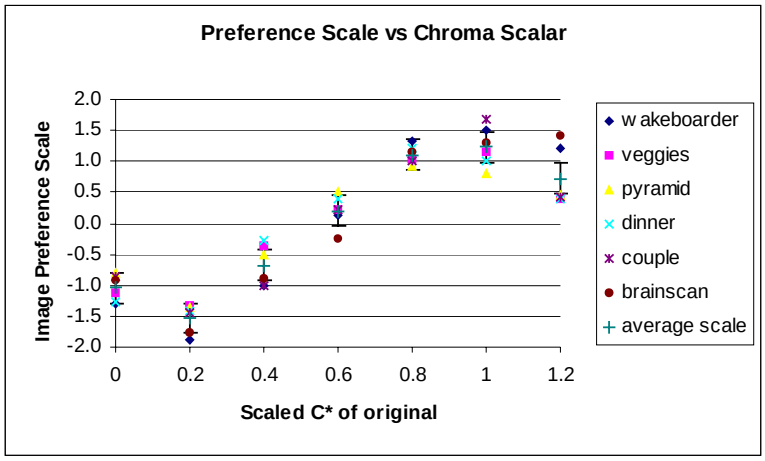


Figure 16. Image preference scale vs. chroma scalar for the six test images.

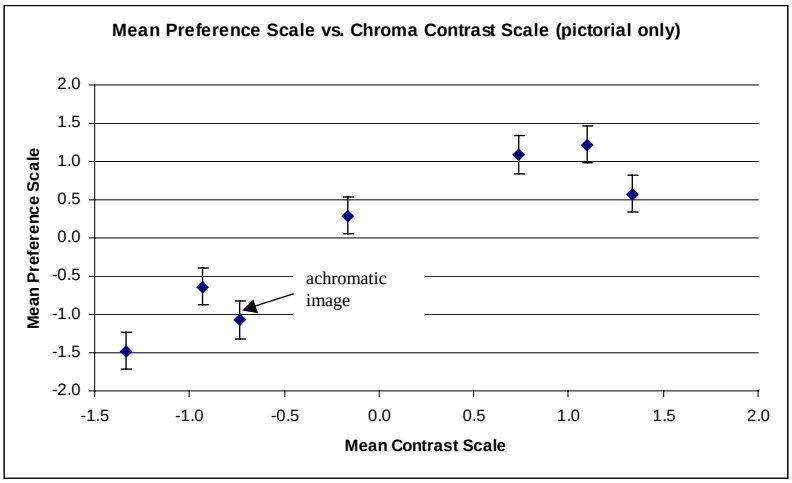


Figure 17. Mean preference scale vs. perceived contrast scale (pictorial only).

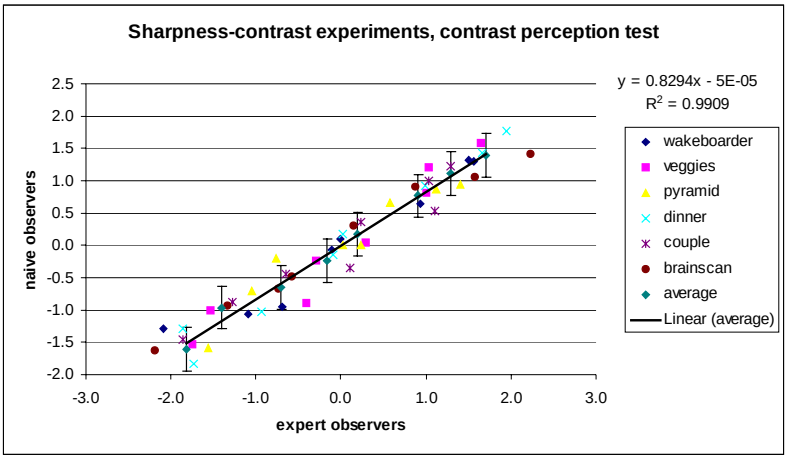


Figure 18. Sharpness-contrast naive observer perceived contrast scale vs. expert observer perceived contrast scale.

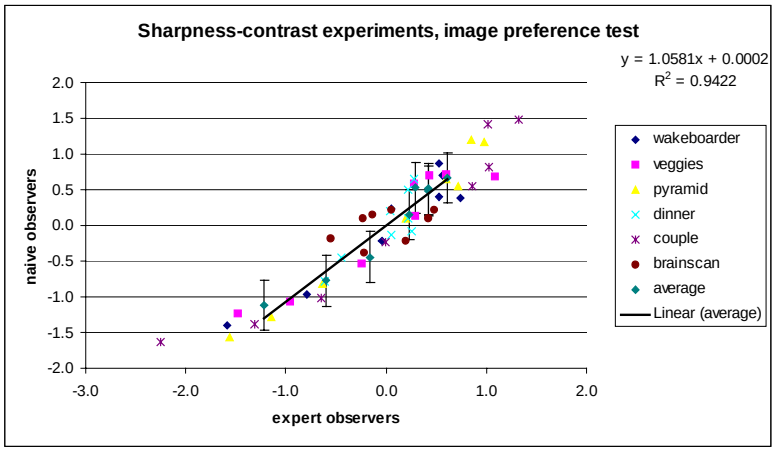


Figure 19. Sharpness-contrast naive observer image preference scale vs. expert observer image preference scale.

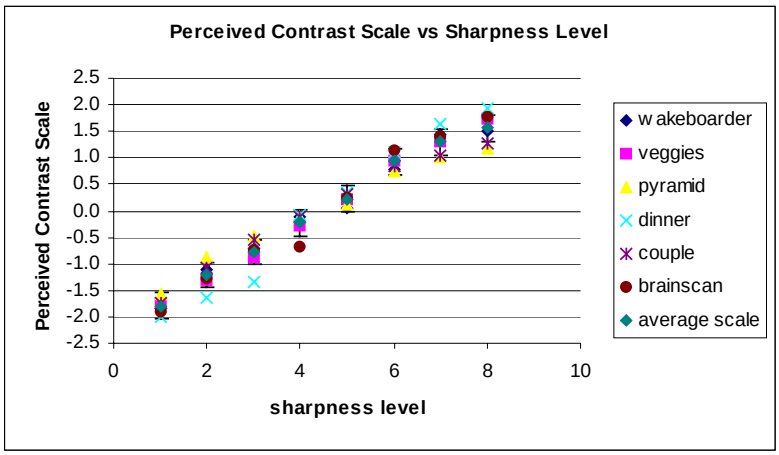


Figure 20. Perceived contrast scale vs. sharpness level for the six test images.

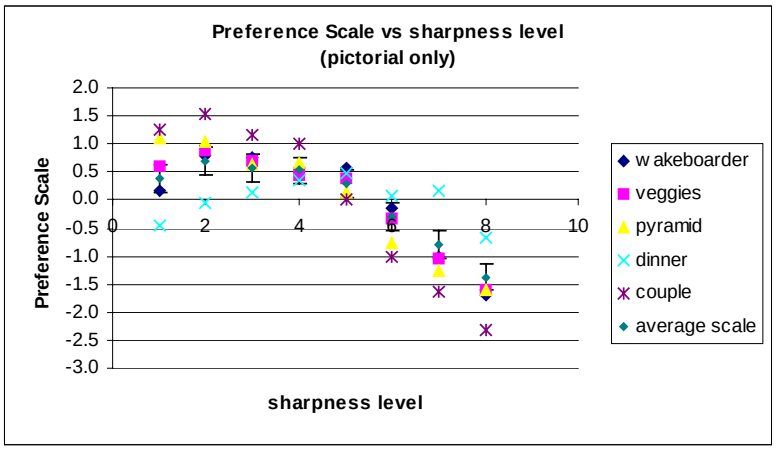


Figure 21. Image preference scale vs. sharpness level for the five pictorial images only (see Table 6-2 for sharpness level).

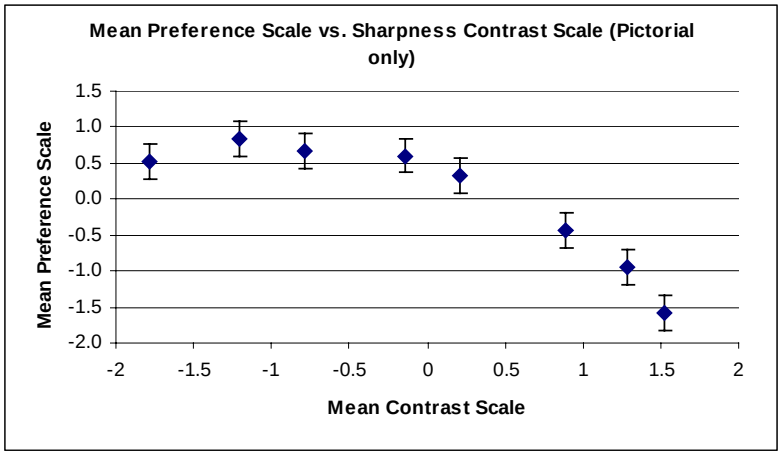


Figure 22. Mean preference scale vs. sharpness-contrast scale for the five pictorial images only.

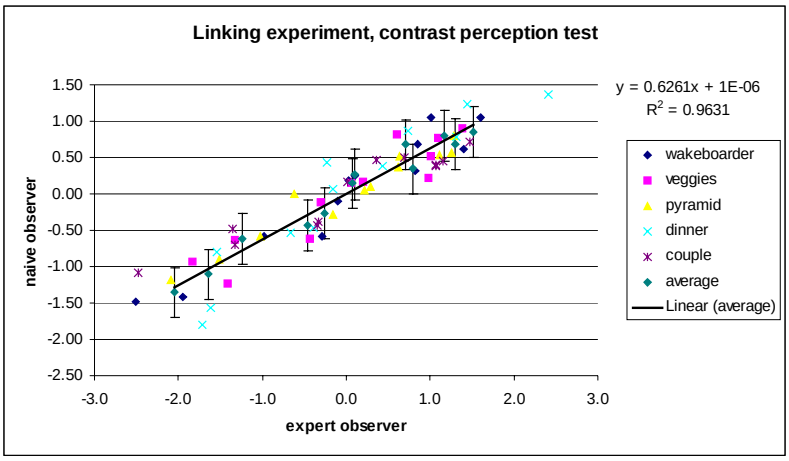


Figure 23. Scale-linking experiment expert observer perceived contrast scale vs. naïve observer perceived contrast scale.

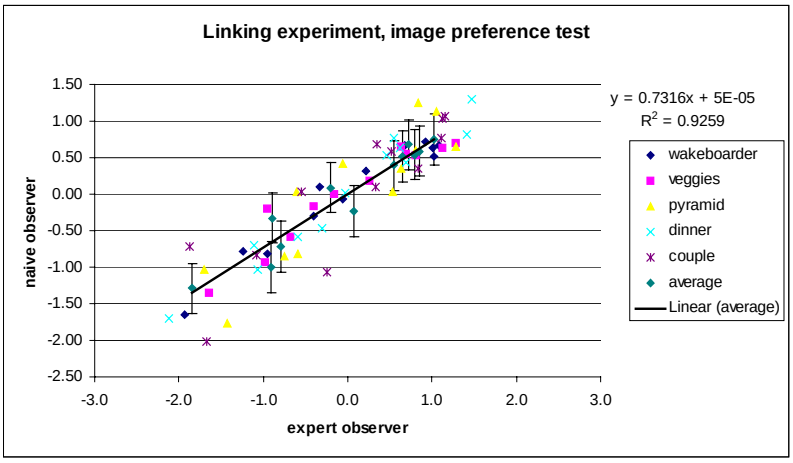


Figure 24. Scale-linking experiment expert observer image preference scale vs. naïve observer image preference scale.

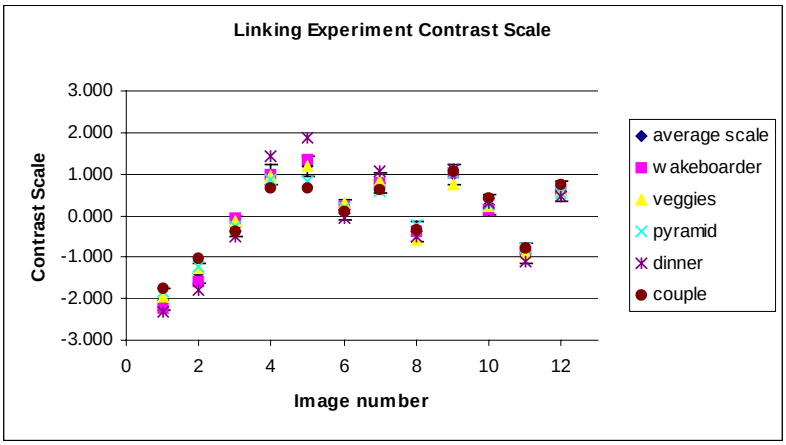


Figure 25. Linking experiment perceived image contrast scale (See Table 4-1 for image numbers).

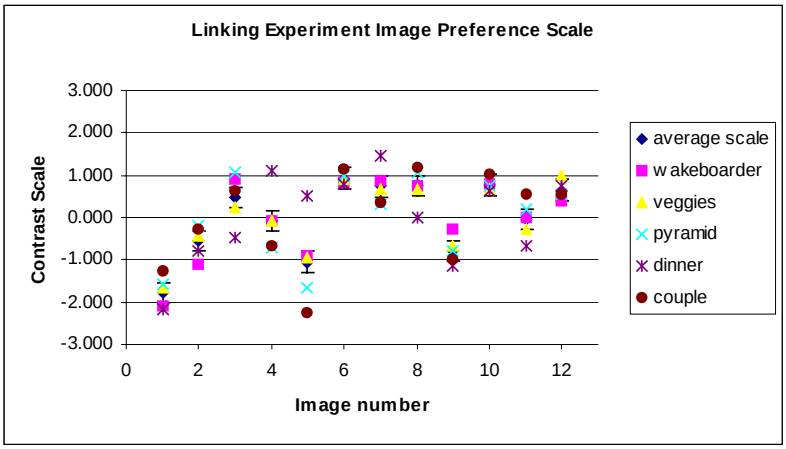


Figure 26. Linking experiment image preference scale (See Table 4-2 for image numbers).

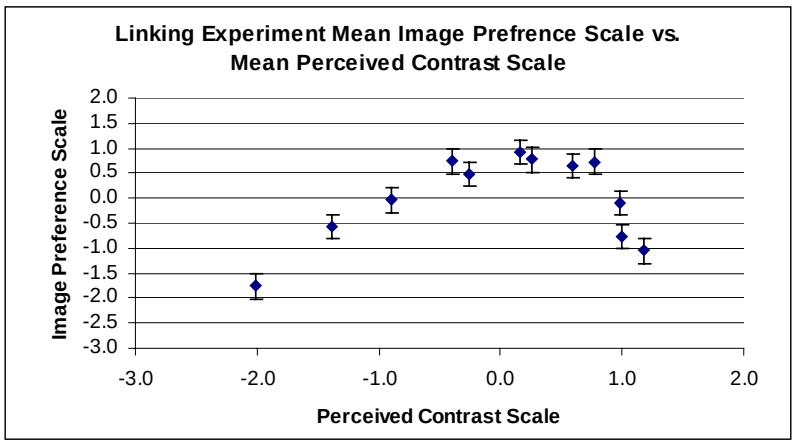


Figure 27. Linking experiment mean perceived contrast scale vs. mean preference scale.

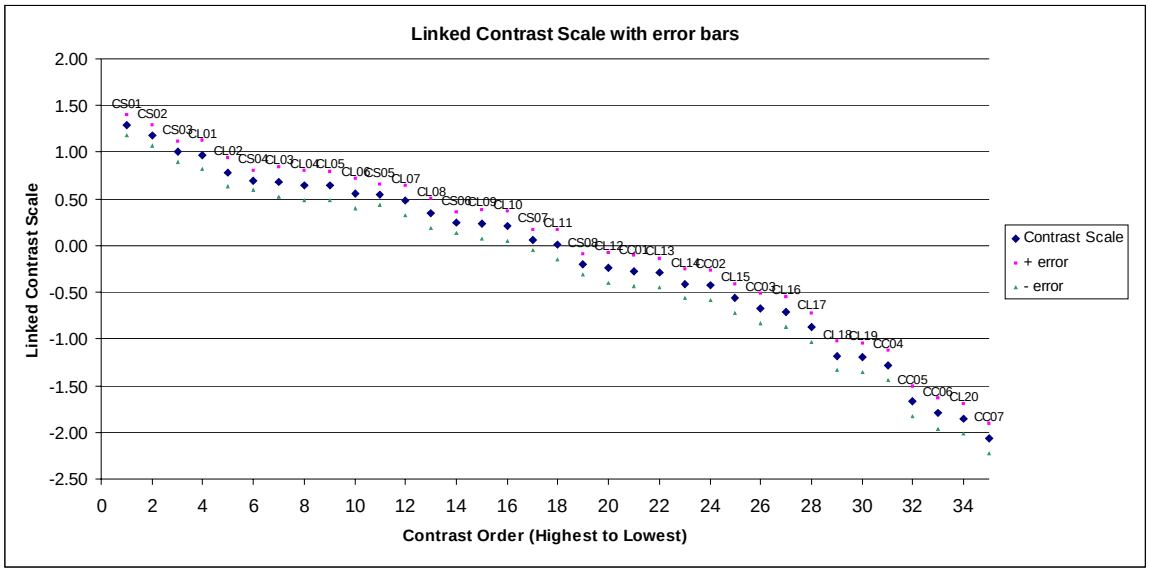


Figure 28. Linked perceived-contrast scale in order from highest perceived contrast to lowest perceived contrast with 95% confidence limits. See Table 7-4 for data labels.

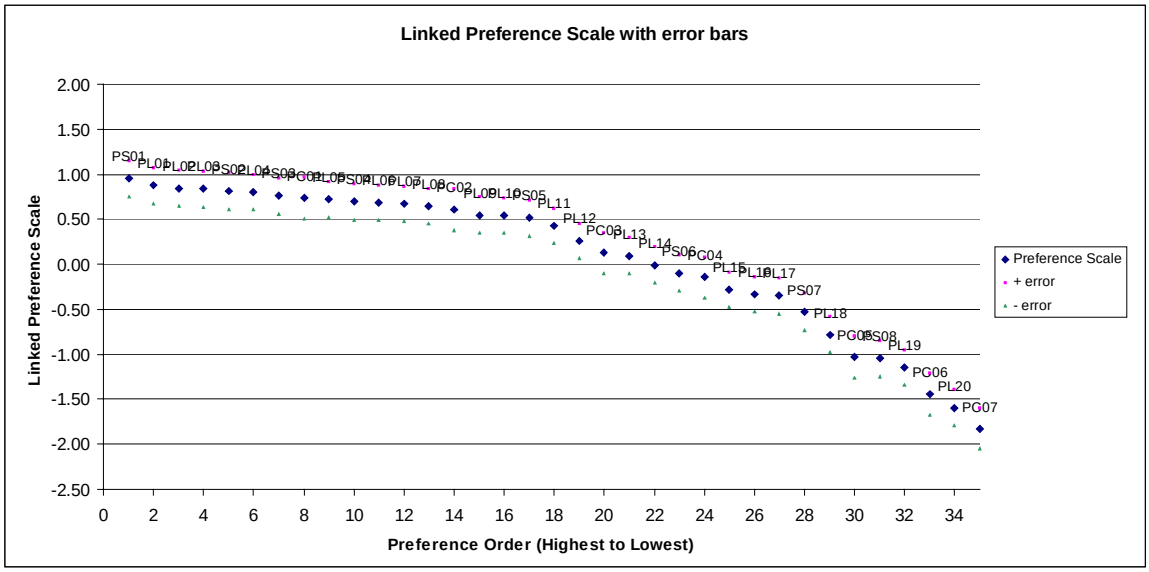


Figure 29. Linked image preference scale in order from most preferred manipulation to least preferred with 95% confidence limits. See Table 7-4 for data labels.

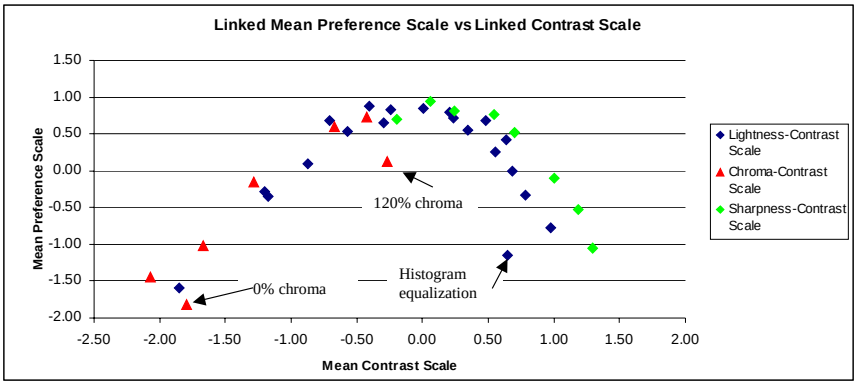


Figure 30. Linked mean preference scale vs. linked mean perceived contrast scale.

Experiment Name	Observer Info Number of Observers	Expert	Naïve	Required Definition of Contrast	Male	Female
Lightness-contrast	32	16	16	8	25	7
Chroma-contrast	32	15	17	10	26	6
Sharpness-contrast	32	15	17	10	26	6
Scale-Linking	32	16	16	9	23	9

Experiment Name	Ethnicity C	A	AA	ME	H
Lightness-contrast	27	3	1	0	1
Chroma-contrast	28	3	1	0	0
Sharpness-contrast	28	3	1	0	0
Scale-Linking	23	6	1	2	0

Table I. Observer statistics from the five experiments. For ethnicity, C=Caucasian, A=Asian, ME=Middle Eastern, H=Hispanic.

L* Manipulation	Manip #	Perceived Contrast Scale (pictorial)	Perceived Contrast Scale (medical)	Preference Scale (pictorial)	Preference Scale (medical)
decrease sigmoid 15	1	16	15	7	9
decrease sigmoid 20	2	14	14	1	18
decrease sigmoid 25	3	12	12	3	5
power 0.90	4	20	20	20	16
power 0.95	5	19	19	15	2
power 1.00	6	11	11	2	15
power 1.05	7	2	3	16	4
histogram equalization	8	3	1	19	13
increase sigmoid 10	9	1	4	18	6
increase sigmoid 15	10	6	7	12	3
increase sigmoid 20	11	8	9	9	1
increase sigmoid 25	12	9	10	5	7
linear 0.05	13	13	13	8	17
linear -0.05	14	10	8	4	11
linear 0.10	15	15	16	10	8
linear -0.10	16	7	5	6	19
linear 0.15	17	17	17	13	20
linear -0.15	18	5	2	11	10
linear 0.20	19	18	18	17	12
linear -0.20	20	4	6	14	14

Table II. Order of perceived contrast and image preference for the pictorial images and the medical image (1 = highest scale value, 20 = lowest scale value).

Contrast Transformations	b0	b1
Lightness-contrast	-0.0877	0.6488
Chroma-contrast	-1.1683	0.6743
Sharpness-contrast	0.6040	0.4513

Preference Transformations	b0	b1
Lightness-contrast	0.1727	0.8060
Chroma-contrast	-0.4246	0.9466
Sharpness-contrast	0.2574	0.8220

Table III. Linear transform parameters used for linking the lightness, chroma, and sharpness-contrast scales.