Color Tolerance Study on White in Practical Aspect: Perceptibility Versus Acceptability

Nooree Na, Kyungah Choi, Jeongmin Lee, Hyeon-Jeong Suk*

Department of Industrial Design, KAIST, #373-1 Guseong-dong, Yuseong-gu, Daejeon, South Korea

Received 9 May 2013; revised 27 July 2013; accepted 30 August 2013

Abstract: This study investigates how the color tolerance for whites is affected by color category and context. Two experiments with different contexts, perceptibility or acceptability, were conducted using 27 color variation samples from six color categories (Neutral-White, Red-White, Yellow-White, Green-White, Blue-White, and Blue-Vivid). The results indicated that the color tolerance for Green-white is dominantly higher relative to the other white-based colors, whereas the color tolerance for Yellow-White and Neutral-White are relatively lower. However, contrary to the hypothesis, no significant differences in the level of color tolerances could be found between the two context of perceptibility and acceptability. The color tolerance thresholds for each color category groups in both experiments were similar.

Col Res Appl, 39, 582–588, 2014; Published Online 7 October 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/col.21847

Key words: white; color tolerance; perceptibility versus acceptability

INTRODUCTION

Ever since McAdam's proposal of the tolerance ellipses plotted on the CIE 1931 chromaticity diagram,[1] many efforts have been made to investigate the phenomena related to color tolerance for human color assessment and consequently, to determine the range of colors that are perceptibly and/or acceptably akin to a given target color. Several researchers have looked into the color tolerances for a range of common colors as well as suggested formulas for calculating their color difference values. Similar to McAdam's research, Melgosa *et al*. discovered that as the values of a^* and b^* (CIE 1976 $L^*a^*b^*$) approach zero, the color tolerance becomes smaller.[2] However, the majority of research that has been conducted primarily focus on evaluating the color tolerances for hue differences of vivid colors. As such, there is a noticeable lack of significant and in-depth studies related to the color tolerances of whites.

White is the standard point of all colors and is one of most abundantly used colors in all industries. Therefore, it seems appropriate that designers and manufacturers both have a fair understanding of white, starting with the fact that the color we refer to as "white" is not a single color, rather a range of colors that are white-based with slight variations in nuances.[3]

With regards to the foregoing statement, this study investigates the differences in color tolerances for five white colors (Neutral-White, Red-White, Yellow-White, Green-White, and Blue-White) and Blue-Vivid using 27 variations for each color. Moreover, previous research is not applicable for the marketing industries in that those experiments were performed without evaluating the colors of real-life, profitable products. Hence, this study also examines how the context of perceptibility, generated by using simple square patch stimuli, and acceptability, generated by smartphone shaped stimuli, affects color tolerance in order to determine the general color tolerance of customers for white-based color products. Considering the fact that white is among the most universally used color in product manufacturing with the greatest level of consumer satisfaction (such as in cars[4] and electronics[5]) due to the influence from the "Apple Effect,"[6] this research will be relevant for not only product designers but also for marketers in producing white colored products that best suit customer needs and preferences.

*Correspondence to: Dr. Hyeon-Jeong Suk (e-mail: h.j.suk@kaist.ac.kr).

RELATED WORKS

With the recommendation of CIE94, interests in color differences have risen significant in the field of color science. Much of the existing research puts an emphasis on color tolerances for hue variations. For example, to quantify the global lack of visual uniformity, Qiao samples hue to investigate the hue discrimination suprathresholds.[7] Complete hue circles at two lightness levels, two Chroma levels, and three given color centers (red, green, and blue) were scaled to produce 39 resulted color centers for hue discrimination.

Taking a further step in the field of color studies, experts have also tried to the determine color difference formulae for predicting the threshold of color differences. In Huang's research, the performance of uniform color space and color-difference formulae for predicting threshold color differences is investigated.[8] The results of the experiment indicate that all formulae and spaces performed similarly to one another and produced relatively consistent calculations with the 25 tolerance ellipses of color found by MacAdam. Similar studies investigate the suprathreshold of ellipsoids by estimating contours of equal color-differences ellipses at 19 color centers. Moreover, with development in new methods of experimentation and technology, the relationships between suprathreshold Chroma tolerances and CIELAB hue angles have also been analyzed to improve the CIE color-difference formulae.[9]

Nowadays, visual assessments of color quality is becoming an influential factor for customers in making a purchasing decision and the principles of color tolerance are becoming more significant in real life situations. Perceptibility and acceptability are two methods of visually assessing color quality that have particularly caught the attention of researchers. Godlove explored the difference in acceptability and perceptibility of dyed fabrics based on the observation that dyed fabric users dislike the changes in hue of fabrics more than the equally perceptible changes in saturation or lightness.[10] In another study, the size of acceptable color differences for wool flannel production was examined, through which it was shown that color difference calculations based on MacAdam's formula correlate best with the visual results.[11]

Similar to the topic of this article, Alghazali *et al*. examines the color tolerances for different shades of white by assessing the perceptibility and acceptability of color differences in artificial denture teeth.[12] According to the experiment results, the mean color perceptibility threshold is significantly lower than the mean color acceptability threshold. However, due to the vast differences in color application and context, it would be inappropriate to employ these findings of color thresholds as pertinent information in the industrial production of consumer goods.

GOAL

The purpose of this study is to understand if color category and context affect color judgment when evaluating variations of color samples and standards that appear to resemble white. Two experiments with six standards and 27 samples for each standard were conducted using the same procedures but with different shaped stimuli to determine whether there was a difference in color tolerance when assessing colors under the context of perceptibility and acceptability. Based on previous studies on color tolerances, the hypotheses are as follows:

H1: There is a difference between the color tolerance thresholds for different nuances of white.

H2: The color tolerance threshold will vary depending on the context of visual assessment (perceptibility or acceptability).

When the context of visual assessment changes, for example by altering the shape or medium of color presentation, users become more conscious about the importance in the difference of color change and process color through different assessment methods. It is hypothesized that product-shaped stimuli, which are assessed by the acceptability method, will have lower thresholds relative to square-shaped stimuli, which are assessed using the perceptibility method.

OBSERVER TEST I: PERCEPTIBILITY

The objective of Observer Test I is to identify the color perceptibility thresholds among colors with high lightness (white-based) and to compare these results with existing studies.

Method

Subjects. A group of 50 people comprised of 25 male students and 25 female students participated in this experiment. The average age of the subjects was 22.20 years with a standard deviation of 2.99 years. In order to confirm that the subjects did not have any color efficiencies, the Ishihara test was completed by all subjects before visual assessment. Moreover, in order to obtain results that are applicable to real-life commercial practices from the point of view of the consumers, all recruited subjects were ordinary people with no educational background in color control.

Stimuli. As stimuli, five colors (Neutral-White, Red-White, Yellow-White, Green-White, and Blue-White) with high lightness ($L^* = 90 \pm 2$) were chosen to represent the white color domain. Blue-Vivid was also selected in order to compare the differences in tolerance for white and vivid colors, mounting to six different color centers, as listed in Table I. Next, well-distributed 27 color samples were made by adding or subtracting 2 from the original R, G, B values of each corresponding center. All stimuli were physically made by printing color patches on white photographic paper with no fluorescence and no gloss. Thus, a total of six standards and 162 samples were prepared for visual examination. The CIE 1976 $L^*a^*b^*$ values for all stimuli were measured using a

TABLE I. Colorimetric specifications of six color centers: CIE 1976 $L^*a^*b^*$ values and L^*c^*h values measured using a spectrophotometer.

Nr.	Color centers	L^*	a^*	b^*	c^*	$h°$
01	Yellow-White	90.52	0.35	5.02	5.03	86.01
02	Red-White	88.86	4.95	0.12	4.95	1.39
03	Blue-White	88.02	−0.02	−5.14	5.14	269.78
04	Green-White	89.83	−5.13	−0.29	5.14	183.24
05	Neutral-White	90.14	0.09	0.08	0.12	41.63
06	Blue-Vivid	41.87	3.39	−30.62	30.81	276.32

spectrophotometer (Minolta CM-2600d). The color differences (ΔE_{ab}) between the samples and their corresponding color centers ranged from 0.22 to 2.77 CIELAB units. Figure 1 shows the distribution of the ΔE_{ab} values. For each of the six color categories, the 27 pairs of printed color patches were made into 65 mm by 65 mm squares for visual assessments.

Procedure. The subjects were given a standard (color center) and a set of samples. For each of the 27 samples, the subjects were asked to assess whether the color appeared "equal" or "not equal" to the standard patch when the two patches were aligned next to each other. The experiment was completed when the subjects evaluated all six sets of samples against the standards. In order to exclude the effects that other colors may have on the appearance of the white stimuli, all color assessments were made with the stimuli placed above a gray color panel with a L^* value of 75. The visual assessments were conducted indoor in front of a large window where a sufficient amount of daylight was received and no artificial light was necessary. The experiments were conducted on clear summer days between 10.00 and 18.00 h, during which there was little variation in color temperature and illuminance. At the subjects' table, the measure of the color rendering index was 87, the correlated color temperature was 6764 K, and the illumination was 835 lx. The visual subtended angle of the stimuli was greater than 4°, which satisfies the CIE reference conditions (Fig. 2).

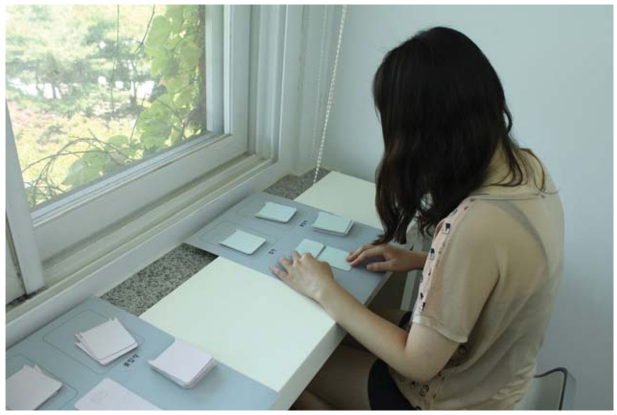

FIG. 2. Environmental settings of visual assessment.

Results and Analysis

The perceptibility levels of all samples were plotted against the color differences for the six color centers. Figures 3 and 4 show the resulting plots for Yellow-White and Blue-Vivid centers.

In every white-based color, there was a tendency for the perceptibility level (how frequently a color sample passes for being perceptibly the same as the color of its standard) to rapidly decline when ΔE_{ab} was greater than 0.40 as shown in Fig. 3. There was a more rapid drop in the average curve to 0 percent perceptibility for samples with large color differences from the standard. Such tendencies in the perceptibility levels against ΔE_{ab} is consistent with Davidson's study.[11] However, for Blue-Vivid, the perceptibility level did not decline to 0 when ΔE_{ab} increased. Rather, the curve showed a flat slope. According to the RIT-Dupont Test,[7] the color discrimination threshold of the CIE-Blue color center was calculated at 3.43. When taking this into consideration as well as the fact that the R, G, B variations of the samples were

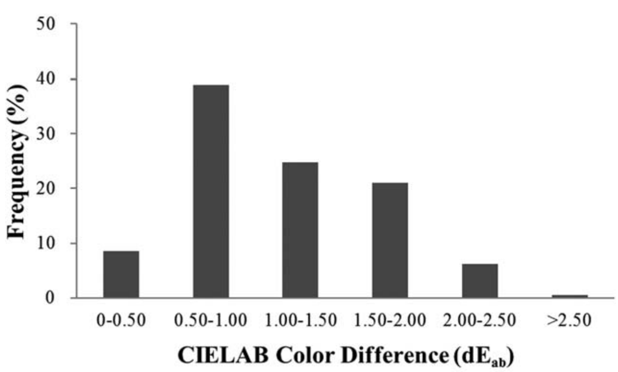

FIG. 1. Distribution of CIELAB dE_{ab} for the 162 stimuli.

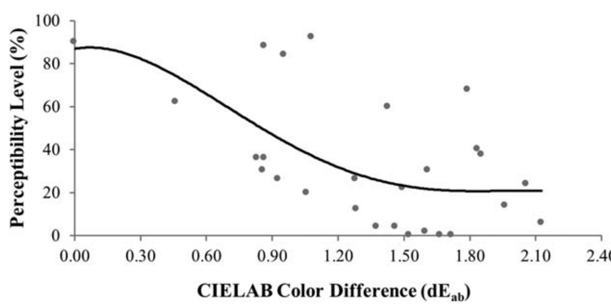

FIG. 3. Percent perceptibility of Yellow-White color samples.

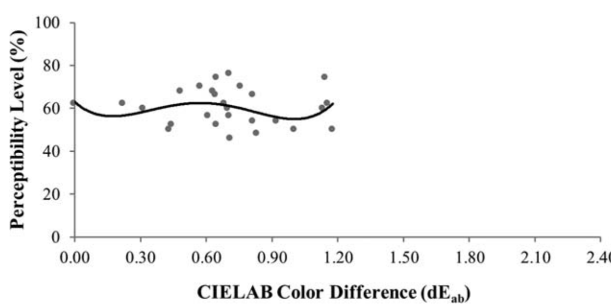

FIG. 4. Percent perceptibility of Blue-Vivid color samples.

TABLE II. Comparison of perceptible chromaticity ellipses at five color centers: semimajor axis (A), ratio of semiaxes (A/B), orientation angle (θ), and the ellipse area (πAB).

Nr.	Color centers	A	A/B	Θ (deg)	πAB
01	Neutral-White	1.11	15.86	90.39	0.24
02	Yellow-White	1.03	14.71	104.27	0.23
03	Red-White	0.75	1.50	33.92	1.18
04	Green-White	1.00	1.35	126.66	2.32
05	Blue-White	0.54	1.74	117.33	0.53

relatively small in this experiment compared with previously conducted researches, the flat curve in Fig. 4 can be seen as just the initial part of the whole curve from the RIT-Dupont Test.

The results were also fitted to draw chromaticity ellipses at the five white color centers in the a^*b^* plane ($L^* = 90 \pm 2$). The standard deviational ellipses were drawn by extracting the samples with perceptibility levels greater than 50%. The resulting ellipses are shown in Fig. 8. The parameters for each ellipse, in terms of semi-major axis (A), ratio of semiaxes (A/B), orientation angle (θ), and the ellipse area (πAB) are given in Table II. As shown, the color tolerance area (πAB) is smallest in the Neutral-White and Yellow-White center, and biggest in the Green-White center. This is consistent with other studies, which indicate that the color tolerance ellipses largely depend on the color category and hue angle.[2,7] According to Xu and Yagu-chi's study, the chromaticity ellipse is biggest for the green center and smallest for the neutral center.[13] Moreover, the orientation angles (θ) of each ellipse except Red-White show that the ellipses are oriented toward the vertical position. Such tendency suggests that human vision is more sensitive to the perception of Red-Green colors rather than Yellow-Blue colors.

OBSERVER TEST II: ACCEPTABILITY

The objective of Observer Test II is to understand the acceptability tolerance of color differences among colors with high lightness and to compare the results with Observer Test I.

Method

Subjects. The subjects from Observer Test I participated in Observer Test II the following day, at the same time of the day under uniform environmental settings.

Stimuli. Staying consistent with Observer Test I, a total of six color centers and 162 samples were prepared for visual assessment. However, the square-shaped stimuli were replaced with smartphone-shaped models (Samsung Galaxy SII) as shown in Fig. 5. The standard colors were made as the upper part of the smart phone, while the 27 samples were made to be the lower part. The stimuli were printed as a 1:1 scale of the actual product (66 mm by 125 mm). Similar to Observer Test I, all stimuli were printed on white nonfluorescent photographic paper with no gloss.

During the experiment, subjects were asked to make a consumer-based decision on color difference by viewing the smartphone-shaped stimuli as real products. This process allowed users to shift from making visual assessments from using the perceptibility method to the acceptability method. As a result of changing the context, the color tolerance levels in Observer Test II were expected to be different from that of Observer Test I.

Procedure. Analogous to Observer Test I, the subjects were given sets of standards and samples and were asked to assess the color difference by joining the lower sample parts of the smartphone to the upper standard part. To evaluate acceptability, the subjects were instructed to evaluate the 27 samples as either a "nondefective product" or a "defective product." The standard for acceptability was based on whether the subject wanted to return the purchased product upon noticing that the two parts of the product were different in color. If the subject decided not to return the product, it was classified as "acceptable" and vice versa. Moreover, the simple Neutral-White square patches were reevaluated in order to prove the validity of Observer Test I. Other environmental settings were controlled to be the same as Observer Test I.

Results and Analysis

For the six color centers, the acceptability levels (how frequently the color differences between a sample and a

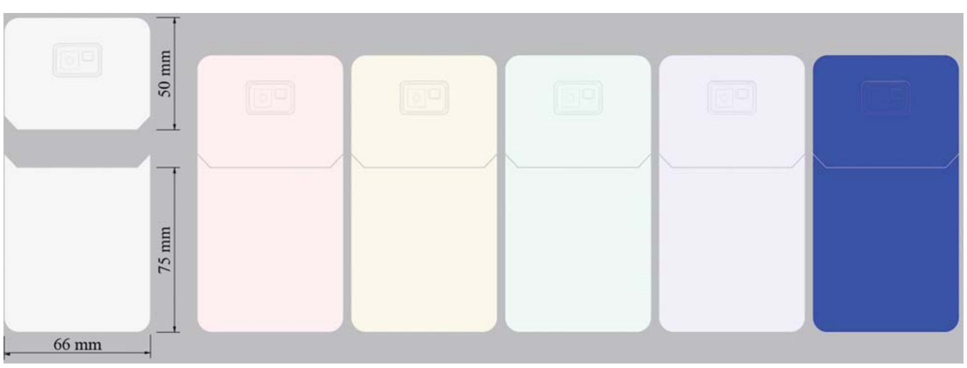

FIG. 5. Smartphone shaped stimuli for acceptability test. Subjects were asked to evaluate the color difference by aligning the top part and the bottom part of the smartphone model. From left to right: Neutral-White, Red-white, Yellow-White, Green-White, Blue-White, and Blue-Vivid.

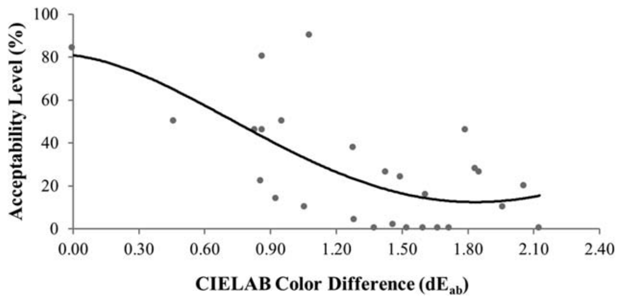

FIG. 6. Percent acceptability of Yellow-White color samples.

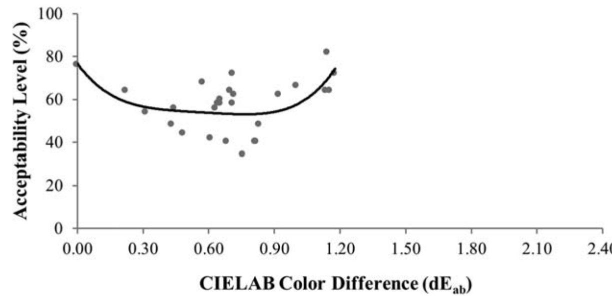

FIG. 7. Percent acceptability of Blue-Vivid color samples.

standard are considered acceptable) for all samples were plotted against the color differences. Figures 6 and 7 show the resulting plots for Yellow-White and Blue-Vivid centers. The average curves are seen to have similar tendencies to the perceptibility curves in Observer Test I.

The results were also fitted as chromaticity ellipses at the five white color centers in the a^*b^* plane ($L^* = 90 \pm 2$). The 5 standard deviational ellipses were drawn by extracting the samples with acceptability levels greater than 50%. The resulting ellipses are shown in Fig. 8, and the parameters for each ellipse are given in Table III. It is shown that the shapes (A/B), orientations (θ), and areas (πAB) of the five color centers are different, which complies with the results in Observer Test I.

PERCEPTIBILITY VERSUS ACCEPTABILITY

Paired T-Test was conducted to determine if the dates on which the experiments were conducted (day of the week) had any significant psychological influence on the subjects' methods of color assessment. The result showed that there was no difference ($p = 0.23$) in the average color tolerance between the 2 days, indicating that the

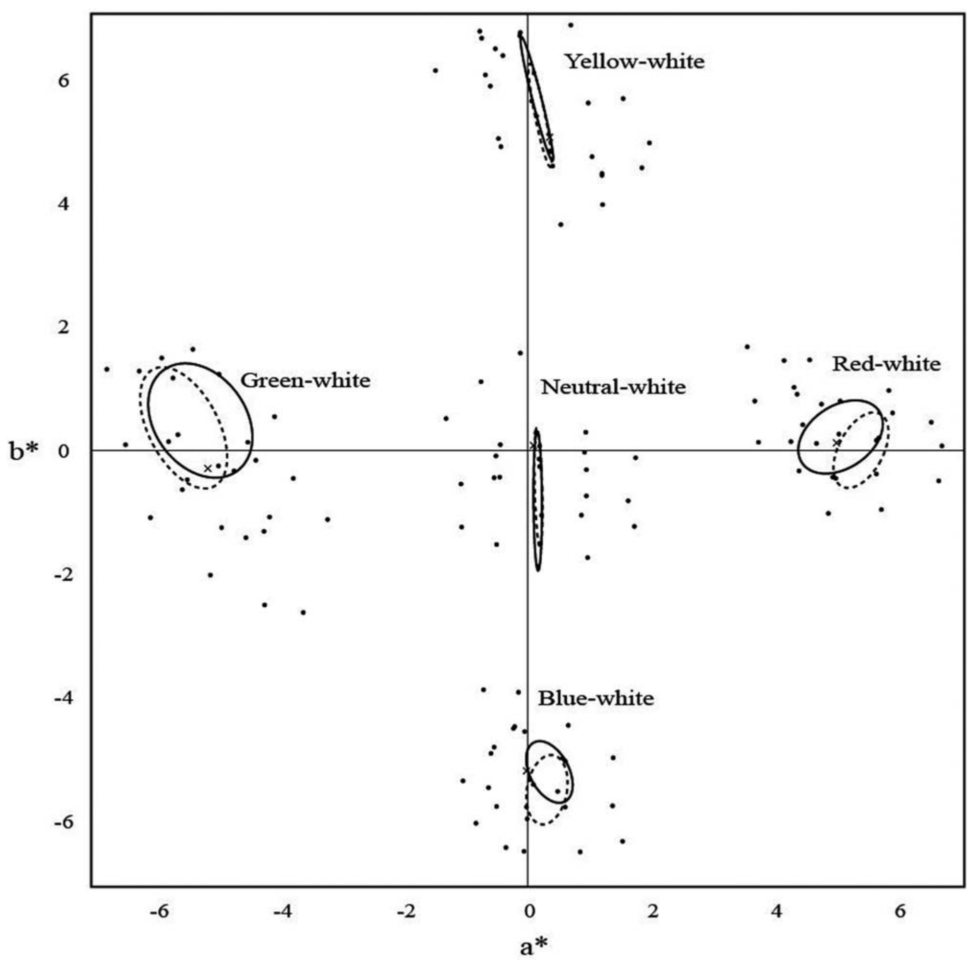

FIG. 8. Visual comparison between the perceptibility ellipses (solid line) and acceptability ellipses (dashed line) in a^*b^* plane for five color centers (cross) with 135 samples (dot).

TABLE III. Comparison of acceptable chromaticity ellipses at five color centers: semimajor axis (A), ratio of semiaxes (A/B), orientation angle (θ), and the ellipse area (πAB).

Nr.	Color centers	A	A/B	Θ (deg)	πAB
01	Neutral-White	0.96	16.00	91.62	0.18
02	Yellow-White	0.85	10.63	102.26	0.21
03	Red-White	0.66	1.83	62.29	0.75
04	Green-White	1.06	1.89	117.42	1.86
05	Blue-White	0.56	1.75	79.87	0.56

days of the week have no effect on how the subjects assess color differences.

In order to compare the color tolerances for perceptibility and acceptability, the chromaticity ellipses were drawn at 5 white color centers as shown in Fig. 8. The solid lines illustrate the perceptibility ellipses, while the dashed lines illustrate the acceptability ellipses. The plotted dots show how 135 samples were distributed around the five color centers (marked with crosses).

In contrast to the hypothesis, the paired T-test indicated that there was no significant difference between the two differently shaped stimuli ($p = 0.81$). It was initially hypothesized that there would be a difference in color tolerance depending on stimuli shape, as such variations in shape establish different contexts and thereby alter the methods of visual assessment (perceptibility and acceptability). However, there were no differences between the color difference levels for perceptibility and acceptability. Godlove's previous study[10] on perceptibility and acceptability also showed that perceptible and acceptable color differences were roughly consistent, contrary to his initial hypothesis. He suggested that significant differences were not expected in some color regions, while in other regions, the difference might be rather large.[14] Table IV shows the color tolerance of perceptibility and acceptability by calculating the method of CIELAB, CIEDE2000, CMC, and CIE94, respectively.

GENERAL DISCUSSION

Various studies have been conducted to show that the color perceptibility threshold is significantly lower than the color acceptability threshold when assessing the color difference for a given set of color stimuli. In contrast to such results, this experiment concluded that there was no significant difference between the tolerance thresholds for the two types of visual assessment.

The insignificance in the differences of the perceptibility tolerance and the acceptability tolerance could have resulted for several reasons. First, it could have been due to the lack of clear experimental instructions. When assessing color difference for acceptability, the context would have been more accurately portrayed if the subjects were instructed to think of themselves as customers trying to purchase a smartphone. However, with the absence of such clear instructions, it appears that many subjects played the role of producers and/or pedantic inspectors during the product manufacturing process rather than customers, creating a perceptibility context instead.

Second, it is possible that the color sample size was too small for evaluating color differences for perceptibility and acceptability. In this experiment, 27 sample swatches were evaluated against one standard. When compared with other experiments, this sample size is relatively small. This could have been the reason as to why there were minimal differences in the tolerance ellipses of product (smartphone) samples and square patch samples, indicating that perceptibility and acceptability judgments were fairly identical. Therefore, it would be meaningful to reconduct this experiment using a more diverse range of color samples with increased color differences to evaluate how color tolerance thresholds vary for perceptibility and acceptability.

Additionally, this article primarily focused on the issue of chromaticity rather than brightness of white. Hence, the brightness values among the color patches should have been consistent and unified. Unfortunately, brightness was not a controlled variable for the sample stimuli, particularly because of the difficulty in producing paper-based stimuli without altering the brightness. This preparatory error could have resulted in inefficient data for visual assessment, especially for evaluating color patches with much higher or lower brightness compared with the standard centers of the stimuli (with CIE L^* values of 90 ± 2).

Finally, the selection of the color region on the CIE diagram could have had an impact on the results. The purpose of this experiment was to examine the color tolerances for different shades of white. Based on data from the RIT-Dupont Test, the color discrimination threshold for gray is $A = 1.33$. But as chroma for gray increases, the threshold increases as well, ranging from to

TABLE IV. Difference of color tolerance between perceptibility versus acceptability.

		Perceptibility				Acceptability			
Nr.	Color centers	CIELAB	CIEDE2000	CMC	CIE94	CIELAB	CIEDE2000	CMC	CIE94
01	Neutral-White	1.10	0.85	1.03	0.99	0.97	0.82	1.00	0.84
02	Yellow-White	0.81	0.67	0.75	0.62	0.81	0.77	0.86	0.72
03	Red-White	0.91	0.91	0.95	0.84	0.80	0.74	0.81	0.73
04	Green-White	0.97	0.90	1.04	0.88	0.58	0.50	0.52	0.53
05	Blue-White	0.65	0.46	0.57	0.65	1.03	0.91	1.37	1.02

2.13 to 3.43. Such results indicate that as chroma becomes lighter and as the shade of color approaches white, color tolerance becomes more sensitive. Hence, another way of looking at this research is as exploring the sensitive levels of color tolerance for different hues of white-based colors by expanding and examining the tolerance area for gray from Haisong's study.[13] However, the focus on such a small area on the CIE diagram could have been the reason as to why the tolerance thresholds for all the white-based colors, particularly Neutral-White, are so small in comparison with other researches on color tolerance, as well as to why there were no significant differences in the tolerance for perceptibility and acceptability.

Although the results of this experiment did not satisfy the initial hypothesis, this study provides opportunity for further research. For example, a supplementary research can be conducted to determine how the type of products affects consumer's level of color tolerance for acceptability. By comparing the acceptability color tolerance of smartphone shaped stimuli with the acceptability color tolerance of stimuli in the shape of other products, it might also be possible to determine how consumer emotions play a role in visual assessment. Such information can be relevant to producers in the manufacturing process, providing them with knowledge on which products to be more assiduous with when dealing with color.

CONCLUSION

This research was aimed at determining whether there were any significance differences in the color tolerance of white-based colors with consideration to the following variables: color category and context. Supporting existing researches, Green-White has the highest color tolerance, whereas Neutral-White and Yellow-White have the lowest color tolerance. However, contrary to the hypothesis of this experiment, no significant differences were found in color tolerance thresholds for when there were changes in the context of visual assessment. Both the color tolerance ellipses of perceptibility and acceptability roughly covered the same areas. In conclusion, it appears that sample variations of ± 2 in RGB values are not sufficient enough for examining how the color tolerance for perceptibility and acceptability differ.

1. MacAdam DL. Visual sensitivities to color differences in daylight. J Opt Soc Am 1942;32:247–273.
2. Melgosa M, Hita E, Poza A, Alman DH, Berns RS. Suprathreshold color-difference ellipsoids for surface colors. Color Res Appl 1997;22: 148–155.
3. Na N, Suk H, Lee J. Investigation of the emotional characteristics of white for designing white based products. Korean J Sci Emotion Sensibility 2012;15:297–306.
4. Schmidt G. DuPont 2011 Global Automotive Color Popularity Report 2011;59.
5. Han SH, Kim KJ, Yun MH, Hong SW, Kim J. Identifying mobile phone features critical to user satisfaction. Hum Factors Ergon Manuf 2004;14:15–29.
6. Elmer-DeWitt P. Apple changed the color of cars, says BMW designer. Fortune 2012.
7. Qiao Y, Berns RS, Reniff L, Montag E. Visual determination of hue suprathreshold color-difference tolerances. Color Res Appl 1998;23: 302–313.
8. Huang M, Liu H, Cui G, Luo MR, Melgosa M. Evaluation of threshold color differences using printed samples. J Opt Soc Am A 2012;29: 883–891.
9. Melgosa M, Huertas R, Yebra A, Pérez M. Are chroma tolerances dependent on hue-angle? Color Res Appl 2004;29:420–427.
10. Godlove I. Improved color-difference formula, with applications to the perceptibility and acceptability of fadings. J Opt Soc Am 1951;41: 760–770.
11. Davidson HR, Friede E. The size of acceptable color differences. J Opt Soc Am 1953;43:581–589.
12. Alghazali N, Burnside G, Moallem M, Smith P, Preston A, Jarad FD. Assessment of perceptibility and acceptability of color difference of denture teeth. J Dent 2012;40 (Suppl 1):e10–e17.
13. Xu H, Yaguchi H. Visual evaluation at scale of threshold to superthreshold color difference. Color Res Appl 2005;30:198–208.
14. Godlove I. Perceptibility and acceptability of color-changes in fastness tests and "on-tone" fading. Am Dyestuff Rep 1951;40:549–558.
15. Berns RS, Alman DH, Reniff L, Snyder GD, Balonon-Rosen MR. Visual determination of suprathreshold color-difference tolerances using probit analysis. Color Res Appl 1991;16:297–316.
16. Cui G, Luo MR, Rigg B, Li W. Colour-difference evaluation using CRT colours. Part I: Data gathering and testing colour difference formulae. Color Res Appl 2001;26:394–402.
17. Cheung M, Rigg B. Colour-difference ellipsoids for five CIE colour centres. Color Res Appl 1986;11:185–195.
18. Guan SS, Ronnier M. Investigation of parametric effects using large colour differences. Color Res Appl 1999;24:356–368.
19. Shen S, Berns RS. Evaluating color difference equation performance incorporating visual uncertainty. Color Res Appl 2009;34:375–390.
20. Berns RS. Billmeyer and Saltzman's Principles of Color Technology. New York: Wiley; 2000.