ACP
AMERICAN COLLEGE OF PROSTHODONTISTS

Color Difference Thresholds of Maxillofacial Skin Replications

Rade D. Paravina, DDS, PhD,[1] Goran Majkic, PhD,[2] Maria del Mar Perez, PhD,[3] & Sudarat Kiat-amnuay, DDS, MS[1]

[1]Assistant Professor, Department of Restorative Dentistry and Biomaterials, The University of Texas Dental Branch at Houston, Houston, TX
[2]Assistant Professor, Texas Center for Superconductivity and Advanced Materials, University of Houston, Houston, TX
[3]Associate Professor, Department of Optics, Faculty of Sciences, University of Granada, Granada, Spain

Keywords
Color; skin; maxillofacial; prosthesis; threshold; visual; spectrophotometer.

Correspondence
Rade D. Paravina, Department of Restorative Dentistry and Biomaterials, The University of Texas Dental Branch at Houston, 6516 M.D. Anderson Blvd., Ste. 493, Houston, TX 77030-3402. E-mail: Rade.Paravina@uth.tmc.edu.

Previously presented at the 24th Annual Houston Conference on Biomedical Engineering Research, Houston, TX, (2007).

A version of this article's abstract was published in the Journal of Esthetic and Restorative Dentistry (2009;21:133–139).

Accepted August 29, 2008

doi: 10.1111/j.1532-849X.2009.00465.x

Abstract

Purpose: The purpose of this study was to determine perceptibility and acceptability thresholds for color differences in light and dark skin-colored maxillofacial elastomers.
Materials and Methods: A total of 15 pairs of light specimens (mimicking white, Asian, and Hispanic skin) and 15 pairs of dark specimens (mimicking African-American skin) were made using skin-colored maxillofacial silicone elastomers, combined with opacifiers and pigments. Color match/mismatch and acceptable/unacceptable mismatch of each pair of specimens were visually evaluated by 45 evaluators under controlled conditions of a viewing booth. Color differences were calculated using CIELAB and CIEDE2000 formulae. After calculating the model parameters, receiver operating characteristics (ROC) curves and area under the ROC curve (AUC) were analyzed. Repeated measures ANOVA and Tukey's HSD test were used in a statistical analysis ($\alpha = 0.05$)
Results: CIELAB/CIEDE2000 perceptibility and acceptability thresholds for light specimens were 1.1/0.7 and 3.0/2.1, respectively. Corresponding values for dark specimens were 1.6/1.2 and 4.4/3.1, respectively. Differences in primary specimen color and type of threshold were found to be significant ($p < 0.001$). Only the primary specimen color effect was found to be significant in AUC comparisons.
Conclusions: Within the limitations of this study, both main effects of threshold type (perceptibility and acceptability) and primary color (light and dark) on 50:50% color-difference thresholds of colored maxillofacial elastomers were found significant for both color-difference formulae used (CIELAB, CIEDE2000). In addition, significant interaction between the two main effects was found, indicating a stronger effect of skin type on acceptability than perceptibility thresholds. Primary specimen color (light vs. dark) was found to be the only significant main effect on the AUC of ROC curves constructed from logistic regression.

Skin is the largest organ of the human body. Its optical properties have attracted much attention in a number of medical disciplines as well as in scientific and industrial disciplines such as scientific imaging, cosmetics, and the stage and screen industries.[1] Maxillofacial prosthetics is a clinical specialty that uses man-made materials to restore function and appearance of missing biologic structures.[2,3] Head and neck defects or deformities can occur due to cancer, trauma, or birth defects.[4-6] Cancer is the most frequent cause of head and neck defects, and people with less pigmentation in their skin are more affected by skin cancer.[4] African–Americans, Asians, and Hispanics are less affected than whites, but are not immune.[5,6] These defects can profoundly affect the patient's quality of life and can impose both emotional and financial burdens for the patient and the family.[7-9] There are numerous reports of dissatisfaction with the esthetics, color stability, function, or longevity of facial prostheses.[10-14] More recently, the result of a global satisfaction quality-of-life test instrument, the "Toronto Outcome Measure for Craniofacial Prosthetics" questionnaire, revealed that the restoration of esthetics was the most relevant need of patients who wear facial prostheses and ranked as the greatest problem among all ten evaluated domains.[15]

Appearance is perception in which the spectral and geometric aspects of a visual stimulus are integrated with its illuminating and viewing environment.[16] Although color is probably the most pronounced appearance attribute of facial skin, the importance of its translucency and gloss should not be underemphasized. Color is a psychophysical sensation provoked in the eye by the visible light and interpreted by the brain. Color dimensions are hue (color name), value (lightness, achromatic, from black to white), and chroma (color strength, from pale to strong). CIELAB system, developed by CIE (Commission Internationale de l'Eclairage, International Commission of Illumination), is a frequently used color notation system.[1] A color-difference formula based on the CIELAB system, CIEDE2000,[17] has recently been introduced and is also used in dental color research.[18] Color difference, which represents the result of all color coordinate differences, is denoted as ΔE^* and $\Delta E'$ in CIELAB and CIEDE2000, respectively.

Perceptibility and acceptability visual judgments are performed in many industries for research, quality control, and related purposes. Perceptibility judgment studies typically consist of subjects answering the question "can I see a difference in color?" These judgments are associated with detecting just-perceptible differences and they do not use interpretation of their importance. These differences are usually small compared to what would be considered acceptable color difference in a given industry.[19] The color-difference tolerance can be determined by asking an additional question, "is this difference in color acceptable?" Therefore, color tolerance for a certain product is the just-perceptible difference increased by a commercial factor.[19] A perfect or almost perfect color match in dentistry can be defined as a color difference below 50:50% perceptibility threshold, which is a color difference that can be detected by 50% of observers (the other 50% of observers will state that there is no difference in color).[20-22] Acceptable match is a color difference below the 50:50% acceptability threshold, which is the color difference described as acceptable by 50% of observers (the other 50% of observers will say that the difference is not acceptable and ask for color corrections or replacement of the restoration).[20-22]

Evaluation of perceptibility and acceptability thresholds is of importance to dental research, just as it is to professional color science.[23-26] For dental applications, thresholds have been mostly analyzed for tooth-colored dental materials, rendering a variety of methods and results;[27-31] however, the data on color-difference thresholds of skin replications are limited to only one paper that evaluated light and dark hand prostheses.[32] Due to the void in the dental literature and differences in location, anatomy, and morphology between the head and neck area and hands, the objective of our study was to evaluate color-difference thresholds of maxillofacial skin replications. To be applicable to the full diversity of patients, it was critical to include multi-racial skin color replications in this study. Light specimens simulated white, Asian, and Hispanic skin colors, while dark specimens simulated African-American skin. All specimens were made using the typical formulations to fabricate facial prostheses for patients of these racial groups in the clinical setting. The null hypothesis was that there was no difference between corresponding color-difference thresholds for light and dark skin replications.

Material and methods

A total of 15 pairs of light specimens (mimicking white, Asian, and Hispanic skin) and 15 pairs of dark specimens (mimicking African-American skin) were made using skin-colored maxillofacial silicone elastomers, combined with opacifiers and pigments. Upon approval from the Committee for the Protection of Human Subjects (no. HSC-DB-04-028), color match/mismatch and acceptable/unacceptable mismatch of each pair of specimens were visually evaluated by 45 evaluators under controlled conditions of a viewing booth. Color differences were calculated using CIELAB and CIEDE2000 formulae. After calculating the model parameters, receiver operating characteristics (ROC) curves and area under the ROC curve (AUC) were analyzed.

Specimen fabrication

Thirty experimental pairs of silicone elastomer specimens and 15 pairs for each light and dark skin shade replications were fabricated by combining an opacifier (titanium white artists' oil color) with a mixture of oil pigments. The elastomer specimens were made from a mixture of a 3:1 ratio of type A medical adhesive to MDX4-4210 medical grade silicone elastomer. Table 1 provides information on the materials used in this study. The opacifier was first mixed with MDX4-4210 to ensure consistency of specimens. Next, the oil pigments were added to the first mixture using a 1.0-ml tuberculin syringe (Becton, Dickinson and Company, Franklin Lakes, NJ). Then, type A medical adhesive (serving as a catalyst) was added and mixed well by hand with a spatula until the color was evenly distributed. Mixtures were then loaded into a syringe (Monoject syringe, Sherwood Medical Co., St. Louis, MO) ready to be injected into the mold.

The specimens ($25 \times 50 \times 3$ mm^3) were processed in metal molds ($90 \times 135 \times 3$ mm^3) with four windows of the same size against two-sided gypsum molds (Die Keen Ivory, Modern Material, Heraeus Kulzer, South Bend, IN) on the top and bottom ($140 \times 190 \times 22$ mm^3). The molds were then placed into a denture flask press (Hanau Engineering Co., Inc., Buffalo, NY) and tightened. The material was allowed to polymerize at room temperature for 24 hours. The specimens were then removed, trimmed, and labeled for identification purposes.

Color assessment and measurement

Visual color assessments were made by 45 color-normal observers—dental professionals (color deficiency was the only exclusion criterion), whose color vision was tested using Ishihara's Charts.[33] A viewing booth (Judge II, GretagMacbeth, New Windsor, NY) with neutral gray walls and floor (Munsell N7 gray) was used for color assessments. Visual color comparisons were performed using D50 illuminant and 0/45° optical geometry. External (overhead) lights were turned off during assessments. Specimens were positioned on the floor of the viewing booth and observed in edge-contact along their longer sides, at a distance of 30 cm and with a visual angle of subtense (2θ) of 10°. The observers rated the color difference as a perfect match, acceptable mismatch, or unacceptable mismatch, with

Table 1 Light and dark skin shade specimen formulations (ml). Pair/Specimen: each pair (1 to 15 for each type of specimen) contained two specimens (1 and 2); for example, 8,10/1 means that the first specimen of the 8th and 10th pairs are made using the formulation given in corresponding row

Pair/specimen	BL	BS	YO	WH	MD	AA
Light skin shade specimens						
1,3,9/1	–	0.02	0.01	0.15	4.00	12.00
1,7,10/2	–	0.02	0.02	0.15	4.00	12.00
2/1; 3/2	–	0.02	0.01	0.15	2.00	6.00
9/1; 2,8,11/2	–	0.02	0.02	0.15	2.00	6.00
4,11,15/1; 6,15/2	–	0.04	0.04	0.15	8.00	24.00
8,10/1; 4,2,12/2	–	0.03	0.01	0.15	2.00	6.00
5,6,13/1; 13/2	–	0.04	0.01	0.15	4.00	12.00
7,14/1; 12,14/2	–	0.02	0.02	0.15	8.00	24.00
Dark skin shade specimens						
1–11/1	–	0.07	0.04	0.11	2.00	6.00
1/2	0.01	0.07	0.04	0.11	4.00	12.00
2/2	0.01	0.07	0.04	0.11	2.00	6.00
12/1, 3,12/2	0.02	0.07	0.04	0.11	2.00	6.00
4/2	–	0.08	0.04	0.11	4.00	12.00
5/2	–	0.08	0.04	0.11	2.00	6.00
13/1; 6,13/2	–	0.10	0.04	0.11	4.00	12.00
7/2	–	0.08	0.05	0.11	8.00	24.00
14/1; 8,14/2	–	0.08	0.05	0.11	4.00	12.00
9/2	–	0.10	0.07	0.11	8.00	24.00
10/2	–	0.08	0.05	0.11	2.00	6.00
11/2	0.01	0.08	0.05	0.11	6.00	18.00
15/1,2	0.01	0.08	0.05	0.11	12.00	36.00

BL = Ivory Black P115 (Grumbacher, Inc., Sanford, Bellwood, IL); BS = Burnt Siena P023 (Grumbacher, Inc., Sanford, Bellwood, IL); PG = Payne's Gray P156 (Grumbacher, Inc., Sanford, Bellwood, IL); WH = Titanium White (soft formula) P250-9 (Grumbacher, Inc., Sanford, Bellwood, IL); YO = Yellow Ochre P244 (Grumbacher, Inc., Sanford, Bellwood, IL); MD = MDX4-4210 (Factor II, Lakeside, AZ); AA = Adhesive Type A (Dow Corning Corporation, Midland, MI).

corresponding grades of 0, 1, and 2, respectively. These grades were then converted to two binary variables: perceptibility (Per) and acceptability (Acc). The value Per = 1 corresponds to observations below perceptibility threshold (perfect match) and Per = 0 to those above the threshold (mismatch). The acceptability variable (Acc) is defined in the same manner as the Per variable, but is defined for an acceptability threshold instead (Acc = 1 corresponding to acceptable mismatch and Acc = 0 corresponding to unacceptable mismatch).

The color of all specimens was additionally evaluated using a spectrophotometer (Color-Eye 7000, GretagMacbeth LLC) set to CIE D50 standard illuminant, 10° observer (CIE 1964 Supplementary Standard Observer), specular component included (SCI), and medium area view (MAV) aperture. Spectral reflectance values in the visible range were recorded in increments of 10 nm. The spectrophotometer was calibrated prior to measurements, in accordance with standard procedure as suggested by the manufacturer. Software (ProPalette 5.0, GretagMacbeth) parameters allowed the data to be presented as the average of a series of three measurements. All results were recorded into a spreadsheet (Microsoft Excel 2000, Microsoft, Redmond, WA).

Data analysis

Perceptibility and acceptability probability against color-difference curves for each observer and experimental condition were calculated using logistic regression.[34] In this study, the proportional odds model was found not appropriate, as it was determined that the shape parameters for the acceptability and perceptibility curves were significantly different. Therefore, a four-parameter logistic regression was performed simultaneously on perceptibility and acceptability scores for each observer and experimental condition. The target function was constructed using the maximum likelihood approach, and the solution was computed numerically via minimization of the negative of the log-likelihood function. The results were processed by a code programmed in software for technical computing (Matlab 7, MathWorks, Natick, MA). The representative values for the perceptibility and acceptability thresholds were then taken as the color-difference values that corresponded to the 50% probability for assigning perfect and unacceptable match grades for the perceptibility and acceptability thresholds, respectively. The procedure was implemented separately for CIELAB and CIEDE2000 color-difference formulae. The CIELAB color differences (ΔE^*) were calculated as follows:[35]

$$\Delta E^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}.$$

The CIEDE2000 color differences ($\Delta E'$) were calculated as follows:

$$\Delta E' = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'_{ab}}{k_C S_C}\right)^2 + \left(\frac{\Delta H'_{ab}}{k_H S_H}\right)^2 + R_T\left(\frac{\Delta C'_{ab}}{k_C S_C}\right)^2\left(\frac{\Delta H'_{ab}}{k_H S_H}\right)^2},$$

where S_L, S_C, and S_H are weighting functions that adjust the total color difference for variation in a perceived magnitude with variation in the location of color-difference pair; k_L, k_C, and k_H are parametric factors (under reference conditions, they are all set at 1).[35]

The perceptibility and acceptability grades from some observers rendered inconsistent values of color-difference thresholds. One common cause of inconsistency was that regression resulted in negative threshold values for some observers who declared imperfect match for one or more specimens with low color difference, while declaring perfect match for some specimens with higher color difference. The second common problem was the case where some observers graded all specimens with the same grade (e.g., imperfect match for all specimens), resulting in a constant p value independent of color difference. To minimize these problems, all observers received initial training and calibration under the same experimental condition. They were asked to compare control pairs of colored specimens that exhibited the smallest color difference (pair number 15 for light specimens, pair number 7 for dark specimens, Table 2), and color difference $\Delta E^* > 5$ (pair number 11 for light specimens, pair number 3 for dark specimens, Table 2). In the vast majority of cases, consensus was achieved on a

Table 2 Color difference (ΔE^* and $\Delta E'$) among evaluated pairs of light and dark specimens

Pair	Light		Dark	
	ΔE^*	$\Delta E'$	ΔE^*	$\Delta E'$
1	1.0	0.7	4.2	2.5
2	2.9	1.6	4.4	2.8
3	2.3	1.2	5.9	4.0
4	3.9	2.5	2.0	1.7
5	2.4	1.9	4.2	2.8
6	3.1	2.3	5.5	3.6
7	2.1	1.5	0.7	0.4
8	4.7	3.4	1.9	1.5
9	2.9	1.5	3.4	2.4
10	4.0	3.0	1.4	1.1
11	5.7	3.9	2.3	1.8
12	4.0	2.9	1.6	1.1
13	0.9	0.6	2.0	1.4
14	1.8	1.3	1.0	0.9
15	0.1	0.1	1.8	1.5

Table 3 Lightness, chroma, and hue angle of light and dark specimens calculated using CIELAB and CIEDE2000 color coordinates: mean, standard deviation (SD), minimum, and maximum values

Color coordinate	Light				Dark			
	Mean	SD	Min	Max	Mean	SD	Min	Max
L*	69.7	2.8	64.9	74.0	48.7	1.2	46.3	50.7
C*	33.3	2.0	28.9	36.7	19.3	2.3	15.0	25.3
h*	54.7	2.0	51.3	57.7	54.6	1.0	55.2	56.4
L′	69.7	2.9	64.8	74.1	48.7	1.2	46.3	50.7
C′	33.8	2.1	29.6	37.3	19.6	2.3	15.2	25.5
h′	53.8	2.2	50.2	57.6	53.7	1.0	51.3	55.2

Light and dark groups consisted of $n = 15$ sample pairs each.

perfect match (for $\Delta E^* < 1$) and unacceptable mismatch (for $\Delta E^* > 5$).

After calculating the model parameters, receiver operating characteristics (ROC) curves were determined from observer grades for all experimental conditions (light–dark, $\Delta E^* - \Delta E'$, perceptibility–acceptability).[36] In addition, ROC curves were also estimated from the logistic model regression data by integrating the probability and 1 − probability against ΔE^* and $\Delta E'$ curves; however, for subsequent data analysis, only the values determined by the former method were used. The area under the ROC curve (AUC) was then calculated to provide a measure of the ability of observers to discern between perfect match/perceptible mismatch and acceptable/unacceptable mismatch as a function of color difference. For the perceptibility threshold example, a theoretical AUC value of 1 for the area under the curve would then denote the perfect situation where there is a single color-difference value (threshold) for which the observer would declare all color differences up to the threshold as a perfect match while declaring a nonperfect match for all values higher than the threshold. For AUC values less than 1, there is no single color-difference value that defines the boundary between perfect match and perceptible mismatch. Instead, there exists a range of threshold values for the perceptibility threshold with different probabilities. The lower the AUC, the less well defined is the color-difference threshold for perceptibility or acceptability. Since the distribution of AUC data was found to be skewed, the AUC data were subsequently ranked.

Repeated measures analysis of variance (RM-ANOVA) was performed on perceptibility/acceptability color-difference thresholds, as well as on AUC values.[37] Because the ΔE values in the two color-difference formulae (ΔE^*, $\Delta E'$) cannot be meaningfully compared, RM-ANOVA for color-difference thresholds was performed for two main factors, primary specimen color or skin type [2-level, light (Lt) and dark (Dk)] and threshold type criterion [2-level perceptibility (Per) and acceptability (Acc)]. For AUC analysis, a third main factor, color-difference formula, was included (2-level $\Delta E^* - \Delta E'$). The main effects and interactions between factors were examined. Pairwise comparisons were performed using Tukey's HSD (Honestly Significantly Different) test. The significance level of $\alpha = 0.05$ was used for all statistical tests.

Results

Color differences (among evaluated pairs of specimens are listed in Table 2. The mean (SD) color difference among light specimens was 2.8 (1.5) and 1.9 (1.1) for CIELAB and CIEDE2000, respectively. Corresponding values for dark specimens were 2.8 (1.7) and 2.0 (1.0), respectively. Mean (SD), minimum, and maximum values of lightness, chroma, and hue angle for two types of specimens, calculated using two color-difference formulae, are listed in Table 3. The difference between mean lightness values of light and dark specimens was $\Delta L^* = \Delta L' = 21$. While light specimens were much more chromatic, the difference between mean chroma values of light and dark specimens was $\Delta C^* = 14$ and $\Delta C' = 14.2$; the corresponding difference in the hue angle was small, $\Delta h^\circ = \Delta h' = 0.1$.

Mean (SD), minimum, and maximum values for perceptibility and acceptability thresholds for light and dark specimens in CIELAB and CIEDE2000 are shown in Table 4. Box

Table 4 Perceptibility (Per) and acceptability (Acc) judgments calculated using ΔE^* and $\Delta E'$ color-difference formula: mean, standard deviation (SD), minimum, and maximum values for perceptibility and acceptability at 50% threshold for light and dark specimens

Formula	Skin type	Threshold	Mean	SD	Min	Max
ΔE^*	Light	Per	1.1	0.5	0.3	2.1
		Acc	3.0	0.4	2.4	4.0
	Dark	Per	1.6	0.5	0.6	2.6
		Acc	4.4	1.0	2.2	5.7
$\Delta E'$	Light	Per	0.7	0.4	0.0	1.3
		Acc	2.1	0.4	1.5	3.0
	Dark	Per	1.2	0.5	0.2	2.0
		Acc	3.1	0.6	2.0	4.3

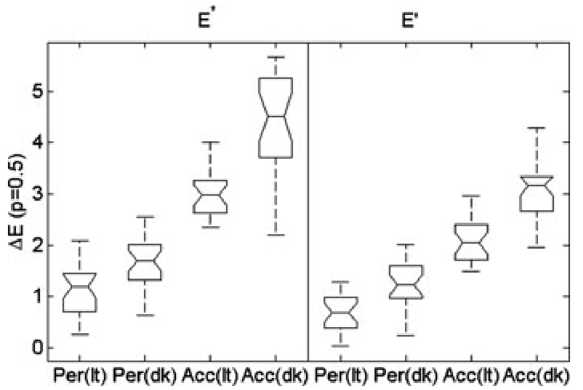

Figure 1 Box plots of perceptibility and acceptability color-difference thresholds for light and dark skin types and two color-difference formulae (left: ΔE^*, right: $\Delta E'$).

plots of acceptability and perceptibility thresholds for light- and dark-colored specimens are given in Figure 1 for both color-difference formulae. An example of perceptibility and acceptability responses (light skin replications, CIEDE2000) for one observer is given in Figure 2. Both main effects (primary specimen color and perceptibility/acceptability threshold) were found significant in ANOVA. In addition, the main effect interaction was also found significant. Tukey's pairwise comparisons for both main effects were also performed for confirmation, finding both main effects significant. To study the interaction, Tukey's pairwise comparisons were performed between each level of the first and second factors. The results revealed significant differences between all levels of the first (second) factor at all levels of the second (first) factor, respectively. An interaction plot between the first and second factor is given in Figure 3, for the CIELAB color-difference formula. The plot reveals a stronger effect of skin type on acceptability threshold than on perceptibility threshold. The same trend is observed for the CIEDE2000 color-difference formula.

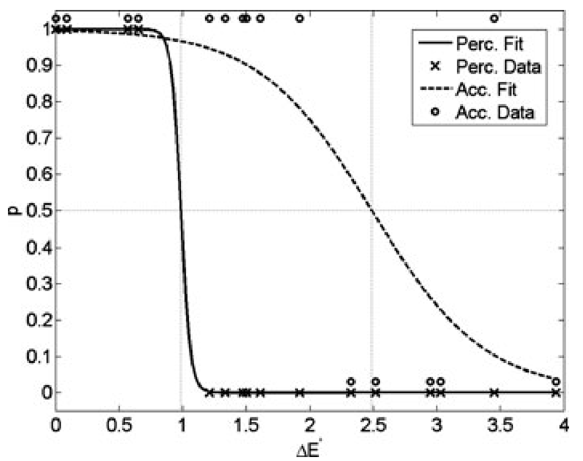

Figure 2 Perceptibility and acceptability responses to light specimens for one observer, light skin replications, CIEDE2000.

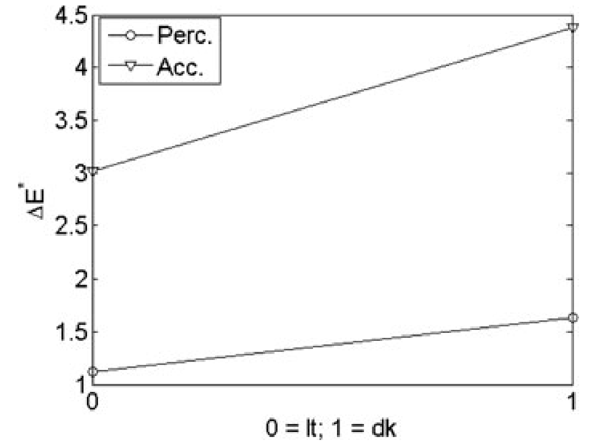

Figure 3 Interaction plot between threshold type (perceptibility—perc, acceptability—acc) and skin type (light—lt, dark—dk) for CIELAB color-difference formula, revealing stronger effect of skin type on acceptability than perceptibility thresholds. The same trend is observed for $\Delta E'$ color-difference formula.

Mean (SD), minimum, and maximum values for AUC for light and dark specimens in two color-difference formulae are shown in Table 5. A box plot of AUC for light and dark skin types, perceptibility, and acceptability thresholds, and two color-difference formulae (left: ΔE^*, right: $\Delta E'$) is shown in Figure 4. ROC perceptibility and acceptability curves (light skin replications, CIEDE2000) for one observer are presented in Figures 5 and 6, respectively. Of all three main effects (primary specimen color, perceptibility/acceptability, color-difference formula), only the first effect (primary specimen color) was found to be significant, while the other two effects (perceptibility/acceptability, color-difference formula) were not significantly different. No significant interaction was found between the three factors; however, the p-value for the interaction between skin type and threshold type (perceptibility, acceptability) was significantly lower than that for other interactions ($p = 0.09$).

Table 5 Perceptibility (Per) and acceptability (Acc) judgments: mean, standard deviation (SD), minimum, and maximum values for AUC for light and dark specimens in ΔE^* and $\Delta E'$ color-difference formula

Formula	Skin type	Threshold	Mean	SD	Min	Max
ΔE^*	Light	Per	0.93	0.09	0.72	1.00
		Acc	0.91	0.09	0.70	1.00
	Dark	Per	0.88	0.08	0.73	1.00
		Acc	0.91	0.07	0.77	1.00
$\Delta E'$	Light	Per	0.94	0.09	0.69	1.00
		Acc	0.92	0.08	0.72	1.00
	Dark	Per	0.90	0.06	0.77	1.00
		Acc	0.93	0.06	0.85	1.00

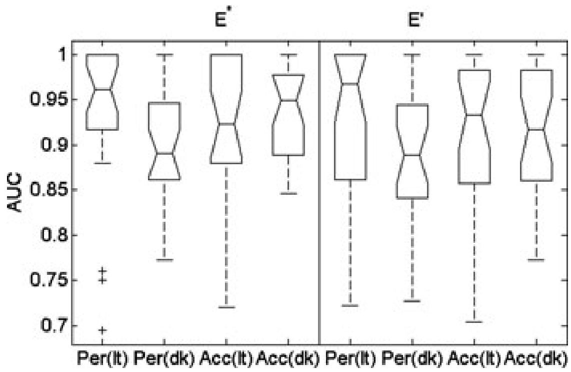

Figure 4 Box plots of AUC for light and dark skin types, perceptibility and acceptability thresholds, and two color-difference formulae (left: ΔE*, right: ΔE′).

Discussion

Mean color differences among pairs of light specimens were almost identical to corresponding values for dark specimens; however, perceptibility and acceptability thresholds for dark specimens were significantly higher than light specimens in both CIELAB and CIEDE2000, thus supporting the rejection of the null hypothesis. Perceptibility and acceptability judgments are identical if color differences are slightly above just-perceivable difference (suprathreshold), and they begin to differ as the color difference among the compared objects increases.[19] The acceptability thresholds recorded in the current study were significantly higher than perceptibility thresholds for each type of skin replication separately and for both color-difference formulae. This finding is logical, as the majority of compared pairs exhibited color difference well above perceptibility threshold.

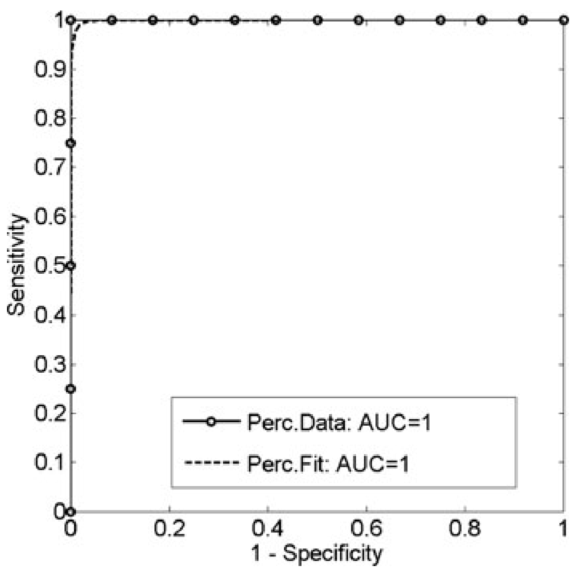

Figure 5 ROC perceptibility curve for one observer, light skin replications, CIEDE2000.

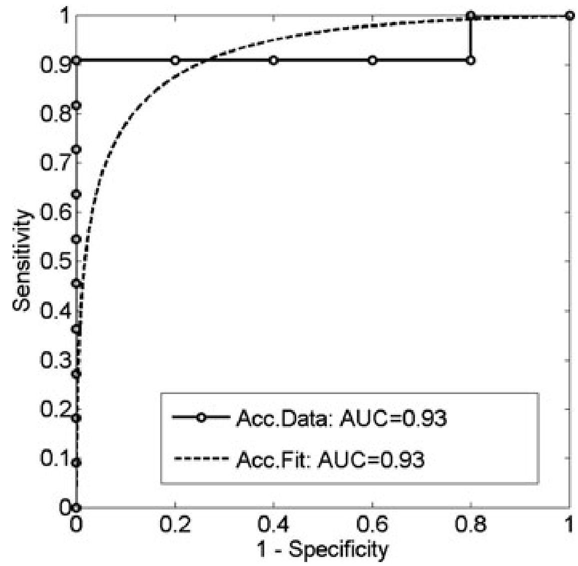

Figure 6 ROC acceptability curve for one observer, light skin replications, CIEDE2000.

The highest single-coordinate difference between light and dark specimens in this study was the lightness difference. Melgosa et al found that visual thresholds for lightness were dependent on the L* value of a sample.[25] Chou et al evaluated 280 pairs of near-neutral matte and glossy paint specimens exhibiting primarily lightness differences.[23] The experimental data were in agreement with the so-called "crispening effect,"[24] defined as the change of perception of color differences caused by the change of background colors. In other words, it was found that gray background increased observers' sensitivity to lightness differences between specimens of the similar lightness as that of the background. In our study, the mean lightness of light skin color replications was much closer to the lightness of walls and floor of the viewing booth than were the dark specimens. Due to a crispening effect, the L* value of the background could have caused lower visual lightness threshold of the light specimens. A study on perceptibility and acceptability judgments of colored surfaces indicated that the chroma discrimination suprathresholds varied, and that less chromatic surfaces exhibited smaller visual chroma thresholds.[25] The current study revealed the opposite, that is, dark skin color replications were more chromatic, and yet exhibited higher thresholds.

Qiao et al reported that the hue discrimination suprathresholds varied with the CIELAB hue-angle position.[26] As the mean hue values in our study were almost identical for both types of skin replications, the differences in the visual hue threshold do not seem to be the cause of the difference in recorded perceptibility and acceptability thresholds for two skin types. Leow et al evaluated color-difference thresholds of two separate sets of ten fair and dark shade digit prostheses fabricated with a stepped increase in color difference as compared to the baseline hand prosthesis.[32] Based on the response from 90 color-normal observers, perceptibility and acceptability color-difference thresholds were $\Delta E^* = 0.8$ and $\Delta E^* = 1.8$ for the fair specimens and $\Delta E^* = 1.3$ and $\Delta E^* = 2.6$ for the dark specimens, respectively.

These values were smaller than reported in Table 4; however, the findings on ethnic-based differences in thresholds exhibited the same pattern: human subjects were less sensitive to color differences in darker shade than fairer shade specimens ($p < 0.001$).

Relatively small skin areas may be considered monochromatic. The same is true for dental materials such as superficial layers of inlay/onlay and restorative materials; however, shade tabs, denture teeth, crowns, bridges, and veneers are polychromatic, as they exhibit color transitions. Several authors reported on the visual thresholds of tooth-colored materials. Wee et al evaluated monochromatic porcelain specimens and reported a perceptibility threshold of $\Delta E' = 1.2$ and an acceptability threshold of $\Delta E' = 1.6$.[27] Douglas et al studied polychromatic denture teeth and reported a perceptibility threshold of $\Delta E^* = 2.6$ and an acceptability threshold of $\Delta E^* = 5.5$.[28] Ragain and Johnston performed their study on monochromatic composite resin disks, and the acceptability threshold was $\Delta E^* = 2.7$.[29] Ruyter et al evaluated monochromatic disks of composite resins for fixed prosthodontic veneers and reported an acceptability threshold of $\Delta E^* = 3.3$.[30] Douglas and Brewer recorded a perceptibility threshold of $\Delta E^* = 0.7$ and an acceptability threshold of $\Delta E^* = 2.1$ for pairs of metal-ceramic crowns.[31]

Variety in obtained results suggests a systematic approach and the standardization of methods for evaluation of color-difference thresholds in dentistry. Future research should address limitations of this study, such as the influence of the experiment setting (in vitro or clinical) and color transitions (monochromatic or polychromatic), and influence of color coordinates on color-difference thresholds.

Conclusions

Within the limitation of this study, both main effects of threshold type (perceptibility and acceptability) and primary color (light and dark) on 50:50% color-difference thresholds of colored maxillofacial elastomers were found significant ($p < 0.001$) for both color-difference formulae used (CIELAB, CIEDE2000). In addition, significant interaction between the two main effects was found, indicating a stronger effect of skin type on acceptability than perceptibility thresholds. Primary specimen color (light vs. dark) was found to be the only significant main effect on the area under curve (AUC) of ROC curves constructed from logistic regression ($p = 0.02$).

References

1. Paravina RD, Powers JM: Esthetic Color Training in Dentistry (ed 1). Philadelphia, PA, Elsevier, 2004, pp. 96-102
2. Cantor R, Webber RL, Stroud L, et al: Methods for evaluating prosthetic facial materials. J Prosthet Dent 1969;21:324-332
3. Andres CJ, Haug SP, Brown DT: Effects of environmental factors on maxillofacial elastomers: Part II–Report of survey. J Prosthot Dent 1992;68:519-522
4. Parkin DM, Pisani P, Ferlay J: Estimates of the worldwide incidence of eighteen major cancers in 1995. Int J Cancer 1993;54:891-903
5. Medina JE, Canfield V: Cancer of the head and neck. In Mayers EN, Suen JY (eds): Malignant Melanoma of the Head and Neck. (ed 3). Philadelphia, Saunders, 1996, pp. 160-162
6. Harrison LB, Sessions RB, Hong WK, et al: Melanoma of the Head and Neck: A Multidisciplinary Approach. Philadelphia, Lippincott Williams & Wilkins, 2003
7. Gotay C, Moore T: Assessing quality of life in head and neck cancer. Qual Life Res 1992;1:5-17
8. Gotay C, Korn E, McCabe M, et al: Quality of life in cancer treatment protocols: research issues in protocol development. J Natl Cancer Inst 1992;84:575-579
9. Gritz ER, Hoffman A: Behavioral and psychological issues in head and neck cancer. In Beumer J III, Curtis TA, Marunick MT (eds): Maxillofacial Rehabilitation: Prosthodontic and Surgical Considerations. St. Louis, MO, Ishiyaku EuroAmerica, 1996, pp. 1-14
10. Kiat-amnuay S, Mekayarajjananonth T, Powers JM, et al: Interactions of pigments and opacifiers on color stability of MDX4–4210/type A maxillofacial elastomers subjected to artificial aging. J Prosthet Dent 2006;95:249-257
11. Lemon JC, Kiat-amnuay S, Gettleman L, et al: Facial prosthetic rehabilitation: preprosthetic surgical techniques and biomaterials. Curr Opin Otolaryngol Head Neck Surg 2005;13:255-262
12. Haug SP, Andres CJ, Moore BK: Color stability and colorant effect on maxillofacial elastomers: Part III. Weathering effect on color. J Prosthet Dent 1999;81:431-438
13. Beatty MW, Mahanna GK, Jia W: Ultraviolet radiation-induced color shifts occurring in oil-pigmented maxillofacial elastomers. J Prosthet Dent 1999;82:441-446
14. Beumer J III, Ma T, Marunick M, et al: Restoration of facial defects: etiology, disability, and rehabilitation. In Beumer J III, Curtis TA, Marunick MT (eds): Maxillofacial Rehabilitation: Prosthodontic and Surgical Considerations. St. Louis, MO, Ishiyaku EuroAmerica, 1996, pp. 377-453
15. Anderson JD, Szalai JP: The Toronto outcome measure for craniofacial prosthetics: a condition-specific quality-of-life instrument. Int J Oral Maxillofac Implants 2003;18: 531-538
16. ASTM: International Standards on Color and Appearance Measurement (ed 7). Conshohocken, PA, ASTM, 2004
17. Luo MR, Cui G, Figg B: The development of the CIE 2000 colour difference formula: CIEDE2000. Col Res Appl 2001;26:340-350
18. Paravina RD, Kimura M, Powers JM: Evaluation of polymerization-dependent changes in color and translucency of resin composites using two formulae. Odontology 2005;93: 46-51
19. Berns RS: Billmayer and Saltzman's Principles of Color Technology (ed 3). New York, Wiley, 2000, pp. 107-131
20. Berns RS, Alman DH, Reniff L, et al: Visual determination of suprathreshold color-difference tolerances using probit analysis. Color Res Appl 1991;16:297-316
21. Ragain JC, Johnston WM: Color acceptance of direct dental restorative materials by human observers. Color Res Appl 2000;25:278-285
22. Lindsey DT, Wee AG: Perceptibility and acceptability of CIELAB color differences in computer-simulated teeth. J Dent 2007;35:593-599
23. Chou W, Lin H, Luo MR, et al: The performance of lightness difference formula. Col Tech 2001;117:19-29
24. Kaneko TA: A reconsideration of Coob-Jubb lightness function. Acta Chromatica 1964;1:103-110
25. Melgosa M, Hita E, Perez MM, et al: Sensitivity differences in chroma, hue, and lightness from several classical threshold datasets. Color Res Appl 1995;20:220-225

26. Qiao Y, Berns RS, Reniff L, et al: Visual determination of hue suprathresholds color-difference tolerances. Color Res Appl 1998;23:302-312
27. Wee AG, Lindsey DT, Shroyer KM, et al: Use of a porcelain color discrimination test to evaluate color difference formulas. J Prosthet Dent 2007;98:101-109
28. Douglas RD, Steinhauer TJ, Wee AG: Intraoral determination of the tolerance of dentists for perceptibility and acceptability of shade mismatch. J Prosthet Dent 2007;97:200-208
29. Ragain JC Jr, Johnston WM: Minimum color differences for discriminating mismatch between composite and tooth color. J Esthet Restor Dent 2001;13:41-48
30. Ruyter IE, Nilner K, Moller B: Color stability of dental composite resin materials for crown and bridge veneers. Dent Mater 1987;3:246-251
31. Douglas RD, Brewer JD: Acceptability of shade differences in metal ceramic crowns. J Prosthet Dent 1998;79:254-260
32. Leow ME, Ow RK, Lee MH, et al: Assessment of colour differences in silicone hand and digit prostheses: perceptible and acceptable thresholds for fair and dark skin shades. Prosthet Orthot Int 2006;30:5-16
33. Ishihara S: Ishihara's Tests for Colour Blindness. London, Hodder Arnold, 1998, pp. 1-24
34. Hosmer DW, Lemeshow S: Applied Logistic Regression (ed 2). New York, Wiley, 2000, pp. 31-203
35. CIE Technical Report: Colorimetry. CIE Pub. No. 15.3. Vienna, Austria, Central Bureau de la CIE, 2004, pp. 12-21
36. Lasko TA, Bhagwat JG, Zou KH, et al: The use of receiver operating characteristic curves in biomedical informatics. J Biomed Inform 2005;38:404-415
37. Crowder MJ, Hand DJ: Analysis of Repeated Measures. Boca Raton, FL, Chapman & Hall/CRC, 1990, pp. 25-70