Perceptibility and acceptability of CIELAB color differences in computer-simulated teeth

Delwin T. Lindsey[a,*] and **Alvin G. Wee**[b,1]

a*Department of Psychology, The Ohio State University, Columbus and Mansfield, OH, USA*

b*Division of Oral Facial Prosthetics, Dental Oncology, Department of Otolaryngology, University of Nebraska, Omaha, NE, USA*

Abstract

Objectives—To determine the perceptibility and acceptability of tooth color differences using computer-generated pairs of teeth with simulated gingival displayed on a calibrated color monitor using appropriate signal detection theory methodology (SDT).

Methods—Twelve dental professionals (four from each of the following groups: dentists, dental auxiliaries, and fixed prosthodontic technicians) and four dental patients served as subjects. Responses to tooth color differences (ΔE) were measured on each of the three principal axes of CIELAB color space ($L*$, $a*$, and $b*$). As a control, responses to $\Delta E = 0$ (the false alarm rate) were also measured in the same experimental session.

Results—No group differences among subjects were found. All gave 50% match or acceptance points that averaged about 1.0 ΔE units in the $L*$ and $a*$ directions, and 2.6 units in the $b*$ direction. False alarm rates across all subjects averaged 27% (4–55%) and 28% (0.4–61%), respectively, for perceptibility and acceptability. A reanalysis of the data based on SDT, which takes subjects' false alarm rates into account, gave somewhat larger color difference thresholds.

Conclusions—Color difference thresholds for our simulated teeth are generally in line with and extend results obtained with studies using "real" dental materials. No differences between thresholds for acceptability versus perceptibility were found.

Furthermore, subjects often reported color differences when none existed, and this behavior needs to be factored into any determination of quality control standards for the fabrication of dental prostheses.

Keywords

Tooth color; Perceptibility; Acceptability; Signal detection; Color difference; Computer simulation

1. Introduction

The human eye is very adept at detecting small color differences between natural and prosthetic teeth. Instrumental color analysis, coupled with advanced fabrication techniques, provides a potentially useful way of minimizing the occurrence of these differences during the fabrication process. However, in order to take full advantage of this technology it is important to establish standards for maximum color differences that are acceptable to dental patients. At present, the

* Corresponding author at: Department of Psychology, 300 Lazenby Hall, Columbus, OH 43210, USA. Tel.: +1 614 292 9535/419 755 4359. E-mail address: lindsey.43@osu.edu (D.T. Lindsey)..
[1]The work was done while he was affiliated with Section of Restorative and Prosthetic Dentistry, College of Dentistry, The Ohio State University, Columbus, OH, USA.

magnitude of perceptible and/or acceptable color differences for human observers is still not well defined in dental color research.

A study by Kuehni and Marcus[1] is cited frequently as a standard for the perceptibility and acceptability of small color differences. A majority of their 63 subjects were experienced, industrial color matchers. They found that the average CIELAB color difference (ΔE CIELAB) for 50% of the observers to perceive a color difference was one unit (ΔE CIELAB = 1). No significant difference between perceptibility and acceptability judgments was found. However, perceptibility and acceptability were assessed using two very different visual tasks, so comparisons may not be valid. Furthermore, Kuehni and Marcus's study tested non-dental materials and spanned a wide color spectrum outside the range of observed tooth colors.

Another study by Seghi et al.[2] evaluated color perceptibility in a group of dental professionals (23 dentists and 4 dental technicians) using translucent color porcelain disks. A visual sorting task, in conjunction with a specialized statistical analysis, was used to estimate perceptibility. Their results suggested that color differences in porcelain disks of about 2 ΔE CIELAB units or greater would be detected 100% of the time and would decline with ΔE. The ΔE CIELAB for 80% probability of observers reporting a color difference was about 0.5–1.0, while ΔE for 50% probability was near zero. Ragain and Johnston[3] also examined color difference acceptability in translucent porcelain disks and reported average acceptability thresholds of 2.72 CIELAB ΔE units. Ruyter et al.[4] who also studied color difference acceptability, found that 50% of their observers (six dentists and six chemists) considered sample pairs of dental composite resin to be unacceptable when ΔE was approximately 3.3 units.

Perceptibility and acceptability were directly compared in a study by Douglas and Brewer.[5] They studied how a group of prosthodontists perceived shade differences in porcelain fused to metal crowns. Acceptability thresholds were measured in the $a*$ and $b*$ directions of CIELAB color space, and were found to be dependent on the direction of color change. Thresholds in the $L*$ direction were not measured. Acceptability of color differences was 1.1 ΔE units in the $a*$ direction and 2.1 ΔE units in the $b*$ direction. Furthermore, they found that while acceptability thresholds (50% point) averaged 2.1 ΔE units, perceptibility thresholds averaged only 0.7 units. However, Douglas and Brewer measured color differences of their translucent samples using instrumentation that exhibited edge loss,[6] so the accuracy of the color difference measurements for perceptibility or acceptability may be significantly in error. Moreover, while Douglas and Brewer[5] attempted to examine the effects of color direction in CIELAB on ΔE, the materials employed in their study did not have color differences precisely confined to any one direction in CIELAB color space.

Finally, Sim et al.[7] had subjects select a Vita Lumin shade tab that best matched each of a series of shade tabs. Selection was made from Vita tabs with a preset color, randomly ordered with respect to color. They then calculated the mean ΔEs between test and matching tabs in the $L*$, $a*$, and $b*$ directions of CIELAB space. The mean perceptible ΔEs estimated in this way were 4.5, 0.69, and 2.4, respectively. Note, however, that these estimates of threshold were not based on independent tests along each color axis, but represented the outcome of subjects attempting to select a match among preset color shade tabs that varied in all three dimensions of color space.

In summary, while there exists a sizeable literature on the perceptibility and/or acceptability of dental color differences, this literature is diverse both with regard to the methodologies employed as well as results obtained in the study of tooth color differences. All the studies that measured the color of actual specimens of dental materials also used non-ideal color measuring instrumentation, which exhibited edge loss.[6]

The present study attempts to address a number of issues raised by these previous studies; specifically, the degrees to which color perceptibility and acceptability (1) differ from one another, (2) vary with direction of change in CIELAB color space, and (3) depend upon the experience of the observer with dental materials. In addition the present study introduces two new methodologies in dental color research: (1) all stimuli are computer-generated and presented under carefully calibrated and reproducible colorimetric conditions and (2) signal detection theory (SDT)[8] is used to evaluate subjects' color difference decisions. SDT has long been used in the field of diagnostic medicine, where there exist important cost/benefit tradeoffs related to the diagnostician's criterion for making diagnostic decisions.[9] In SDT, estimates of perceptibility and acceptability will take into account both the hit and false alarm rates, which to our knowledge have not been used in dentistry to estimate performance on a color visual task.

The objective of this study was to evaluate the perceptibility and acceptability thresholds of dental professionals and patients for tooth color differences along each of the three principal directions in CIELAB color space: *L**, *a**, and *b**. In addition we reanalyzed the date using STD.[8] The following hypotheses were tested in this study:

- *H1*: perceptibility and acceptability thresholds will differ.

- *H2*: perceptibility and acceptability will differ across groups of subjects in the dental profession.

- *H3*: perceptibility and acceptability will differ across different axes in CIELAB color space.

2. Materials and methods

2.1. Subjects

Sixteen subjects participated in this study, four each drawn from the following categories in dentistry: (1) dentists, whose clinical practice was primarily restorative dentistry, (2) individuals of varying dental experience, primarily dental auxiliaries, (3) dental fixed prosthodontic technicians, and (4) dental patients. Subjects were recruited from faculty and staff from the University Medical Center and surrounding city. All had normal color vision as assessed by the American Optical Hardy–Rand–Rittler (AO-HRR)[10] pseudoisochromatic plate test, and all had normal visual acuity with corrective lenses. The study protocol was approved by the University's Institutional Review Board approval (IRB #2003H0019, dated 2 October 2003), and all subjects gave informed consent prior to their participation in the study.

2.2. Apparatus

Measurements of subjects' tooth color difference discrimination and crown acceptance/ rejection were based on computer-generated stimuli presented on a high-resolution ($1600h \times 1200v \times 80$ frames/s $\times$ 8 bit per color channel) 21″ RGB monitor (24D Diamondtron, Mitsubishi/Gateway, North Sioux City, SD). To standardize the subject's viewing distance, subjects sat in a chair with their head supported by a headrest and viewed the monitor screen from a distance of 1.2 m. Subjects used a two-button computer mouse to initiate the presentation of each stimulus and to signal their response: match/no-match or accept/reject.

2.3. Stimuli

Control of stimulus presentation as well as subjects' responses was under the control of custom computer software. Prior to the start of each trial, subjects viewed a uniform gray screen (mean luminance = 31 cd/m^2) with an approximate perceived reflectance of about 0.25. A small black square, located approximately in the center of the intended stimulus region, provided a spot

for subjects to fixate their gaze prior to the presentation of the stimulus. When the subject initiated a trial, a highly schematic depiction of reddish upper gingiva appeared, along with two simulated teeth (approximately $0:5°h \times 0:7°v$) with their cervical portions embedded in the gingiva (Fig. 1). Tooth size, as seen by the subject, corresponded to viewing natural teeth at approximately arm's length. The simulated tooth on the left was always the standard natural tooth of fixed color, and the tooth on the right was the crown, which was to be compared to the natural tooth. The crown differed in color from the standard tooth on exactly one-half of 1500 trials presented in a single experimental session (3 principal directions of color change $\times$ 5 stimuli per direction $\times$ 50 trials per stimulus condition = 750 trials). These trials were randomly interleaved with trials in which standard and crowned teeth were colorimetrically identical. The stimulus remained on the screen for no more than 2 s; it was then replaced by the uniform gray screen. The subject could respond anytime during or after the presentation interval. Doing so terminated the trial and produced a uniform gray screen and fixation dot, signaling that a new trial could be initiated by the subject. The relatively brief presentation time not only meant that a large number of trials could be run in a single experimental session, but also limited the effects of chromatic adaptation on judgments of perceptibility and acceptability. To further limit these effects, the location of the fixation dot varied slightly from trial to trial and subjects were instructed to maintain fixation from trial onset to termination.

The colors of standard and crowned teeth were determined by specifying the R, G, and B values of the display pixels that constituted the teeth. In addition to the eight bits of color resolution in R, G, and B channels available from the graphics card of our computer, spatiotemporal dithering was utilized to obtain two additional bits of color resolution in each of the channels. Spatial dithering was obtained by manipulating the RGB values of adjacent horizontal scan lines defining the teeth. At the viewing distance of the stimuli, the human eye cannot resolve the individual scan lines, and the effective color is determined by the average color of the scan lines. Similarly, temporal dithering was obtained by manipulating the RGB values of the teeth in successive frames of the video image. The eye cannot resolve the small variation in color across successive frames, and the effective color is determined by the average color across successive frames. It was estimated that with this technique, and taking gamma correction into account, the graphics system was, in principle, able to reproduce desired tooth CIELAB values to within 0.15 ΔE.

The colors of the standard natural tooth and various crowns were specified in CIELAB coordinates, relative to a 100% reflectance standard with a correlated color temperature approximating D65. The CIELAB coordinates for the standard tooth were $L* = 78.03$, $a* = 1.35$, and $b* = 22.68$. These coordinates are close to measurements of average central incisor color obtained in a diverse population of patients ($L* = 77.28$, $a* = 4.16$, and $b* = 19.47$).[11]

Subjects made perceptibility and acceptance judgments for three series of five pair of tooth colors. Each series consisted of variation in tooth color confined to one of the three dimensions of CIELAB space. In the $L*$ and $a*$-series, crown tooth color varied approximately in increments of 0.8 ΔE units from the standard tooth. In the $b*$-series, this increment was approximately 1.4 units.

2.4. Calibration

The R, G, and B channels of the monitor were corrected for nonlinearities in the variation in light output with change in pixel value (gamma correction) using standard techniques.[12] Spectral calibration of the R, G, and B phosphors was performed with a SpectraScan PR705 (PhotoResearch Inc., Chatsworth, CA) spectroradiometer placed at the same location in the apparatus as a subject's eye. The CIE 1931 chromaticity coordinates (x, y) and maximum luminances (in cd/m^2) of the three monitor phosphors were R: 0.622, 0.341, 29.37; G: 0.289, 0.604, 107.3; B: 0.151, 0.067, 12.83. These values, along with the gamma corrections were

used to transform the desired CIELAB values for the standard and crown teeth into the appropriate pixel RGB values in the graphics display.[13] Spectral calibrations of the phosphors were performed weekly and were found to vary less than 1% over the time course of the study. Full spectral calibration of the colors of the standard tooth and the 15 simulated crowns used in this study was performed immediately prior to and immediately after final data collection. Little difference was found in the estimates of tooth color obtained at these two times.

2.5. Procedure

Data on each subject were collected in two sessions occurring on different days. Tooth color perceptibility was always tested in the first session, and crown acceptance/rejection in the second session. Prior to making perceptibility judgments, subjects were instructed to base their judgments on criteria given in the following script:

> "We want to know whether or not you think that the two teeth are the same or different color. This means: do they match, are made from the same material, are they identical, etc."

Prior to making crown acceptance judgments, subjects were instructed to base their judgments on criteria given in the following script:

> "We want to know whether or not, in your mouth, you would want a crown the color on the right to be placed next to a tooth the color on the left. This means: the tooth on the left is your real, natural tooth and the one on the right is the tooth that is the possible color for a replacement and we want to know if the possible crown would be an acceptable match."

Subjects sat in a darkened room, illuminated by the light from the computer monitor. Subjects adapted for 10 min to the uniform gray background of the monitor before data collection began. Subjects pressed the left mouse button to initiate a trial. Subjects then pressed the left button to indicate a match (accept) or the right button to indicate no match (reject). Subjects were instructed to rest at any time they wished during the test session. In any event, every 150 trials the computer program instructed subjects to rest before continuing.

2.6. Statistical analysis

The parameter d' derived from STD was used to specify visual sensitivity in a way that took into account both a subject's correct reports of a color difference (hit rate) and his incorrect reports of color differences when none occurred (false alarm rate). For this particular task, d' is usually calculated by

$$d' = z(\text{H}) - z(\text{FA})$$

(1)

where $z(x)$ is the z-score function, and H (hits) and FA (false alarm) are, respectively, hit and false alarm rates.[14] The resulting data were analyzed using a three way ANOVA [dependent measure (d'); two observation types (perceptibility, acceptance) × four subject types (technicians, dentists, patients, auxiliary) × (three color directions (L^*, a^*, and b^*)].

3. Results

Values of percent "no match" and "unacceptable" are plotted in Fig. 2 as functions of CIELAB color difference, ΔE, for our 16 dental professional or patient subjects. Separate results for our four categories of subjects are not shown because there were no statistically significant differences in performances among the categories of observers tested in this study. Separate panels show results obtained for three series of color differences: L^*-, a^*-, and b^*-series. Subjects' responses to no color difference (Color Difference = 0) are also shown in each panel.

The solid line function drawn through the data indicates the average results obtained from the 16 subjects.

There are three noteworthy features of the data. First, within the same series, perceptibility and acceptability tasks gave indistinguishable results. Second, for each task, $L*$- and $a*$-series stimuli gave nearly identical results, while the $b*$-series function rises more gradually as color difference is increased. And third, on average, subjects were quite willing to report no match (average: 27%; range: 4–55%) or unacceptable (average: 28%; range: 0.4–61%) even when there were no color differences between the standard and crown teeth. Note, in particular, that for some subjects, this occurred on more than 50% of such trials.

The horizontal line in each panel shown in Fig. 2 indicates the 50% no match or unacceptable rates. If this response level is used to establish a threshold for perceptible or unacceptable color differences, as has often been the case in prior studies of tooth color difference, then the threshold ΔE values are approximately 1.0, 1.0 and 2.6 for the $L*$-, $a*$-, and $b*$-series, respectively.

The ANOVA showed that color difference was the only significant factor ($p < 10^{-7}$) affecting subject response. No interactions among factors were significant ($p > 0.79$). Bonferroni-corrected paired-comparisons revealed no significant differences between the $L*$- and $a*$-series data, but a significant difference ($p < 0.01$) between $L*$- and $b*$-series, and between $a*$- and $b*$-series.

Fig. 3 shows the perceptibility data re-plotted in terms of d', as calculated using each subjects' hit and false alarm rates. Results for each of the three series of color differences – $L*$-, $a*$-, and $b*$-series – are shown in separate panels, as in the previous figure. The horizontal line drawn at a value of $d' = 1.0$ in each panel corresponds to a level of performance that is often used to specify threshold performance in a variety of psychophysical studies of human vision. For the 27% average false alarm rate found for the perceptibility task, a d' of 1.0 corresponds to an average hit rate of 65%. This level of performance was obtained for ΔE values of 1.25, 1.25, and 2.8 units on the $L*$, $a*$, and $b*$ axes, respectively. Analysis of the acceptability data gave nearly identical results and they are not shown in Fig. 3.

4. Discussion

4.1. Color differences

The results of this study suggest that when estimates of sensitivities to color differences are based on hit rates, average thresholds are found in the range of 1.0–2.8 ΔE units in CIELAB space. These values are in the range previously reported by many others. A difference can also be found in sensitivity with respect to the axes in CIELAB color space: perceptibility and acceptability thresholds based on 50% hit rates were 1.0 in the $a*$ and $L*$ directions, but 2.6 in the $b*$ direction. When thresholds are based on d', ΔE values increase to 1.25 units in the $L*$ and $a*$ directions and 2.8 in the $b*$ direction. These findings are consistent with and extend those of Douglas and Brewer,[5] who compared thresholds in the CIELAB $a*$ and $b*$ directions only.

The dependence of color sensitivity on direction of color change in CIELAB is not peculiar to the conditions of this experiment and occurs because the formulas for converting CIE 1931 CIE colorimetric values to CIELAB $L*a*b*$ coordinates does not adequately capture differences in human visual discrimination for colors approximating human teeth. When results are expressed in CIE 1931 units, it can be found that they are very similar those obtained in a number of classical studies of human color discrimination.[15-18] Ragain and Johnston[3] found that the CMC (1:1) color space provided a better fit to their data than did CIELAB. However,

it is important to recognize that either color space, CMC (1:1) or CIELAB, should work well for representing tooth color differences, once suitable standards for scaling tooth color perceptibility and/or acceptability in different directions of color space have been established empirically.

4.2. Perceptibility versus acceptability

The results of this study also show that perceptible color differences were also unacceptable, regardless of whether judgments were made by dental patients or trained dental professionals. There are few published studies that have directly compared perceptibility and acceptability with which results can be compared. The study by Kuehni and Marcus,[1] which found equivalent threshold ΔEs for perceptibility and acceptability, employed large color patches falling well outside the color range of teeth. More recently, a study by Douglas and Brewer, [19] employing porcelain fused to metal crowns, found that ΔE for acceptability was significantly greater than perceptibility. However, there are a number of important differences between this study and that by Douglas and Brewer[19] that may account for this discrepancy in results. Subjects in the present study could view samples for no more than 2 s, while Douglas and Brewer[19] limited viewing to 10 s. Thus, it might be argued that the brief stimulus exposures used in the present study did not allow enough time for the highly subjective judgment of acceptability to reliably emerge, and subjects settled on any perceptible difference in color as their criterion for acceptability. However, it is also possible that the instructions to subjects encouraged a very conservative criterion for acceptance, as compared to Douglas and Brewer. [19] Subjects in the present study were explicitly asked to judge acceptability on the basis of whether they would want a crown of a given color difference cemented in their mouth. Douglas and Brewer,[19] on the other hand, merely asked subjects whether they would "accept or reject the shade match under clinical conditions". It is also noteworthy that in the present study, subjects made perceptibility and acceptability judgments independently, in separate experimental sessions conducted on different days. In Douglas and Brewer's study,[19] perceptibility and acceptability judgments were not made independently: subjects were asked on a trial by trial basis (1) whether they detected a color difference, and, only if they said "yes", were they asked (2) whether the shade difference was acceptable.

4.3. False alarms

A striking feature of the results is the number of false alarms obtained from subjects. False alarm rate did not depend on the classification of the subject (dental professional or patient), nor did it depend on the task: equal false alarm rates were obtained, on average, on the perceptibility or acceptability tasks. Also, considerable individual variability was found in false alarm rate, as indicated by the psychometric functions shown in Fig. 2. It could be thought that the observed false alarm rates indicate that subjects responded differently to computer generated stimuli than they would have responded to more natural stimuli. This explanation is unlikely since false alarm rates comparable to those found in the present study are evident in the results of others who have used more natural stimuli, even though those authors did not specifically analyze the false alarm rates. For example, examining Douglas and Brewer,[19] who used linear regression to estimate thresholds for overall perceptibility and acceptability, shows false alarm rates of 10% and 40%, respectively. Ragain and Johnson,[3] show acceptability results for two subjects. False alarm rates based on analysis of a figure (see Fig. 8 in Ref. 3) from their published study are estimated to be about 15%.

What factors can account for the false alarm rates found in this study? According to SDT, a subject's decision to respond in a particular way (e.g., match versus no match in the perceptibility task) depends not only upon the sensory signals evoked by standard and comparison stimuli, but also upon the subject's criterion for deciding if the sensory signals arise from the same color or from different colors. A criterion is necessary, according to SDT,

because the responses to visual stimuli (the "signal") are noisy[8] and fluctuate randomly from one trial to the next, even when the same stimuli are presented. Thus, a difference in sensory signals can be evoked by standard and comparison stimuli even when the two stimuli are identical in color. The subject's criterion determines how large a difference between sensory signals must be present before he will say "no match" or, by extension to acceptability judgments, "not acceptable".

The implications of these results, according to SDT, are that a subject's perceptibility or acceptability will depend not only on his sensitivity to small color differences, but will also depend on the context within which the subject is asked to make a judgment. For example, if a subject perceives the need to maximize his performance (hit rate) when small differences between standard and comparison occur, he is likely to adopt a conservative criterion. SDT predicts that this criterion will indeed guarantee a higher hit rate than would be obtained by adopting a more liberal criterion, but his conservative criterion also guarantees a higher false alarm rate, all other things being equal. The graphs of d' shown in Fig. 3, according to SDT, show performances in terms of sensitivity (d') that factor out each subject's response criterion.

A subject's response criterion may depend on many factors, including the perceived demands of the task (Is this a test?) and his prior expectations of the outcome of his decisions. These characteristics will, of course, vary from individual to individual, and may be quite different in a purely clinical setting as compared to the setting in which other investigators have measured perceptibility and acceptability psychophysically. Thus, the results suggest that in order to achieve useful standards for tooth color difference, future studies of color perceptibility and acceptability need to address questions regarding how patients as well as dental professionals make their decisions in the most clinically relevant setting. In any event, a subject's performance on a discrimination task depends upon both his sensory capabilities and his response criterion. SDT allows one to separate these two effects and to study them independently.

The results of the present study were obtained using very crude approximations to human dentition. However, the means exist today to extend the approach for future research to the creation of far more realistic computer simulations of human dentition. These techniques need not be confined to the study of small differences in tooth color. Instead, they can be used to study the perception of small variations in tooth translucency, pigmentation, and surface texture, as well as the interactions of light with these tooth characteristics that occur in the mouth of the simulated patient. The ability to realistically simulate human dentition, coupled with the computer's inherent capacity to automate data collection in an efficient and highly repeatable manner, offers a methodology that promises rapid progress in establishing reliable standards for the fabrication of prosthetic teeth.

5. Conclusions

Computer-generated pairs of tooth-shaped patches of color embedded in stimulated gingival and displayed on a calibrated, computer-controlled color CRT provided an effective way of measuring human sensitivities to tooth color differences. Analysis based on signal detection theory provided an effective way of controlling for a subject's bias in reporting color differences. The results indicated (1) ΔEs of about 1.25, 1.25, 2.8, respectively, in the $L*$, $a*$, and $b*$ directions of CIELAB color space, (2) no significant differences between perceptibility and acceptability, and (3) no significant differences in performance across different groups of subjects selected from the field of dentistry. The methodology employed in this study can be extended easily to the study of color differences in more realistic computer-generated scenarios.

References

1. Kuehni RG, Marcus RT. An experiment in visual scaling of small color differences. Color Research and Application 1979;4:83–91.

2. Seghi RR, Hewlett ER, Kim J. Visual and instrument colorimetric assessments of small color differences on translucent dental porcelain. Journal of Dental Research 1989;68:1760–4. [PubMed: 2600257]

3. Ragain JC, Johnston WM. Color acceptance of direct dental restorative materials by human observers. Color Research and Application 2000;25:278–385.

4. Ruyter IE, Niler K, Moller B. Color stability of dental composite resin materials for crowns and bridge veneers. Dental Materials 1987;3:246–51. [PubMed: 3479360]

5. Douglas RD, Brewer JD. Acceptability of shade differences in metal ceramic crowns. Journal of Prosthetic Dentistry 1998;79:254–60. [PubMed: 9553875]

6. Bolt RA, Tenbosch JJ, Coops JC. Influence of window size in small-window color measurement, particularly of teeth. Physics in Medicine and Biology 1994;39:1133–42. [PubMed: 15552102]

7. Sim CP, Yap AU, Teo J. Color perception among different dental personnel. Operative Dentistry 2001;26:435–9. [PubMed: 11551006]

8. Green, DM.; Swets, JA. Signal detection theory and psychophysics. New York: John Wiley; 1966.

9. Starr SJ, Metz CE, Lusted LB, Goodenough DJ. Visual detection and localization of radiographic images. Radiology 1975;116:533–8. [PubMed: 1153755]

10. Hardy, LH.; Rand, G.; Rittler, MC. AO H–R–R pseudoisochromatic plates. New York: American Optical Company; 1957.

11. Gozalo-Diaz, DJ. Masters Thesis, College of Dentistry. Columbus, OH: The Ohio State University; 2004. Relationship between the color of teeth, gingiva, lips and skin.

12. Lindsey DT, Teller DY. Motion at isoluminance: discrimination/detection ratios for moving isoluminant gratings. Vision Research 1990;30:1751–61. [PubMed: 2288088]

13. Wyszecki, G.; Stiles, WS. Color science: concepts and methods, quantitative data and formulae. New York: Wiley & Sons; 1992.

14. Macmillan, NA.; Creelman, DC. Detection theory: a user's guide. Mahwah, NJ: Lawrence Erlbaum; 2004.

15. MacAdam DL. Visual sensitivities to color differences in daylight. Journal of the Optical Society of America 1942;32:247.

16. Brown WRL, Macadam DL. Visual sensitivity to combined chromaticity and luminance differences. Journal of the Optical Society of America 1957;39:808. [PubMed: 18142394]

17. Brown WRL. Color discrimination in twelve observers. Journal of the Optical Society of America 1957;47:137. [PubMed: 13406654]

18. Wyszecki G, Fielder GH. New color-matching ellipses. Journal of the Optical Society of America 1971;61

19. Douglas RD, Brewer JD. Variability of porcelain color reproduction by commercial laboratories. Journal of Prosthetic Dentistry 2003;90:339–46. [PubMed: 14564288]
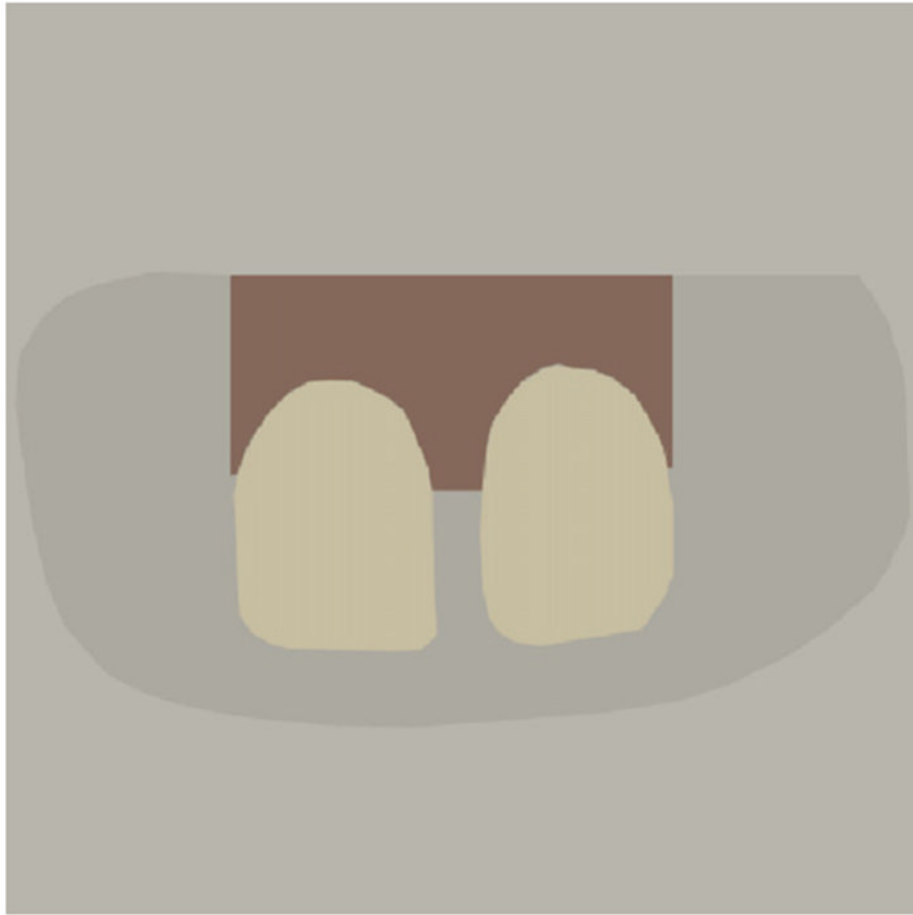
Fig. 1.
Stimulus configuration used to measure tooth color difference perceptibility and crown acceptance/rejection.
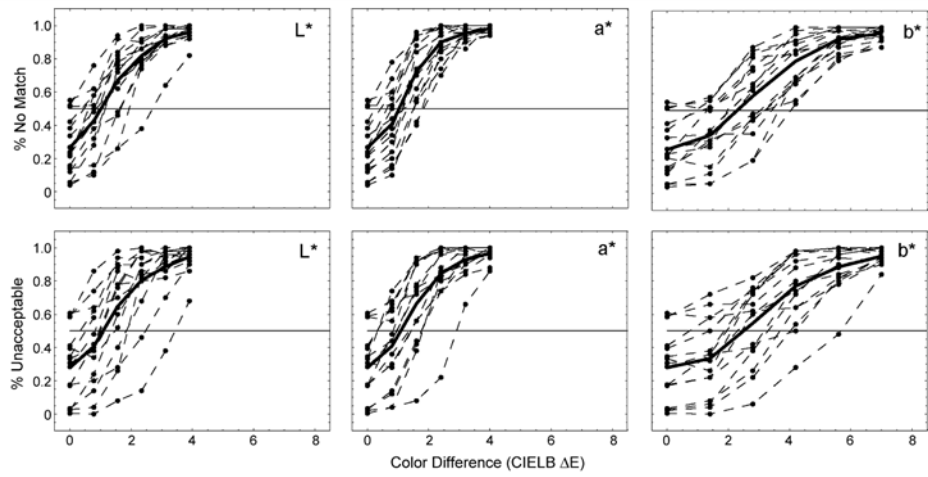
Fig. 2.
Perceptibility (top three panels) and acceptability judgments (lower three panels) as functions of ΔE varying along each of three principal axes of CIELAB color space.
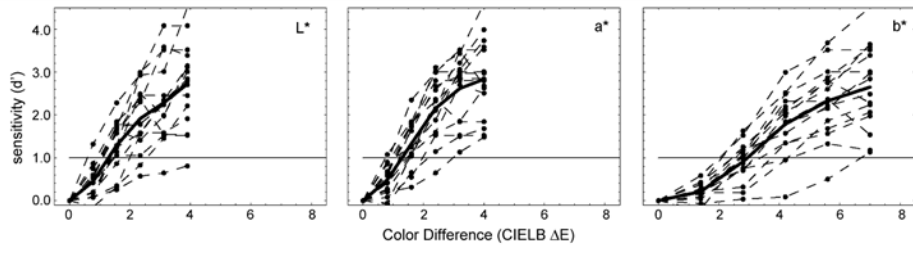
Fig. 3.
Replot of perceptibility data shown in Fig. 2, showing values of d′, as calculated from Eq. (1).