Faculty versus Student Repeatability on Evaluating Translucency of the Anterior Dentition

James L. Sheets 1, David B. Marx 2, Nina Ariani 3, Valentim A. R. Barão 4 and Alvin G. Wee 5,*

1 Department of Prosthodontics, School of Dentistry, Creighton University, Omaha, NE 68178, USA; jls63680@creighton.edu
2 Department of Statistics, University of Nebraska, Lincoln, NE 68588, USA; david.marx@unl.edu
3 Department of Prosthodontics, Faculty of Dentistry, Universitas Indonesia, Jakarta 10430, Indonesia; nina.ariani02@ui.ac.id
4 Department of Prosthodontics and Periodontology, Piracicaba Dental School, University of Campinas (UNICAMP), Piracicaba 13414-903, SP, Brazil; vbarao@unicamp.br
5 Division of Prosthodontics, Department of Restorative Dentistry, University of Minnesota, Minneapolis, MN 55455, USA
* Correspondence: wee00008@umn.edu

Abstract: The objective was to compare the repeatability between dental faculty, whose clinical practice was primarily restorative dentistry, and final year dental students in categorizing the inherent translucency of images selected at random using either a 3- or 7-point scale (translucent to opaque). Digital images of anterior dentition were randomly selected based on inherent translucency. Thirty images (five were repeated) were randomized and categorized by 20 dental students and 20 faculty on their inherent translucency. Statistical analysis was performed using an F test for analysis of variance at 95% confidence interval. A covariance parameter estimate (CPE) was accomplished to compare the inter-rater variability of the dental faculty and dental students. Statistically, more variability occurred between Slides (CPE of 0.185 ($p = 0.001$)) and between Subject and Slide (CPE of 0.122 ($p = 0.0002$)) than within subjects (CPE of 0.021 ($p = 0.083$)). Viewing repeat Slides, Students (CPE = 0.16) were more consistent ($p < 0.05$) than faculty (CPE = 1.8) using the 3-point scale, while the CPE was the same (CPE = 0.669) using 7-point scale. Dental students and faculty were consistent using the 7-point scale to judge repeat slides, while dental students in this limited pilot study were more consistent when viewing a repeat slide using the 3-point scale.

Keywords: prosthodontics; ceramics; color science; laboratory technology

1. Introduction

An individual’s smile is an important factor in their perception of self [1]. If a restoration is planned, communication between the dentist and the dental laboratory is paramount to an esthetic restoration that meets all the specifications required to satisfy the dentist and patient [2,3]. If the clinician cannot describe or illustrate what he or she clinically views during the shade matching process [4], then it becomes difficult for the laboratory technician to fabricate an esthetic restoration. As noted by Fondriest: beyond talent, the ultimate likelihood of faithfully recreating nature is limited by how complete the communication process is [4].

Traditionally, when describing dental esthetics; hue, value, and chroma are emphasized as primary selection factors. Secondary optical properties of the tooth exist and affect the overall appearance of the tooth [1]. These commonly include translucency, opacity, iridescence, surface gloss, and fluorescence [5]. Translucency and opacity are rated as the most important of these secondary properties, since they are an indication of the quality and quantity of light reflectance [6].

Translucency is defined as the property of a substance that permits the passage of light, but also disperses light [7]. The inherent translucency of anterior crowns becomes...
an important aspect of esthetics to ensure a match to the adjacent dentition. According to Spink et al., “the goal of dental ceramics is to imitate a tooth’s color and vitality by recreating an appropriate mix of light absorption and scattering.” [8]. Kelly et al. noted that substrate translucency therefore becomes one of the primary factors in controlling esthetics and is a critical consideration in the selection of materials [9].

Traditionally, translucency evaluation is visually accomplished with the aid of shade guides provided by multiple dental porcelain manufacturing companies. Instrumental measurements of translucent restorative materials using a spectrophotometer, colorimeter, spectroradiometer and digital camera and software have been reported [10,11]. Multiple indices have been used to calibrate translucency: total or direct transmission coefficient (tc), translucency parameter (TP) and contrast ratio (CR) [11].

Liu et al. [10] evaluated the relationship between instrumental measurements and subjective visual assessment of the differences in dental porcelain translucency. Three observer groups (final year dental students, residents, and faculty) took part in the study [10]. Increased shade matching experience ($\geq 10$ years, faculty) significantly improved the ability to perceive differences in translucency [10].

In order to ensure a good inherent translucency match information must be transferred to the dental laboratory technician. Conventionally this is done with a written work authorization form which can be augmented with digital or conventional photographs or internet communication. There is no concise method of categorizing the inherent translucency of the anterior dentition and transferring that information from dentist to dental laboratory technician. A lack of research on the validity or repeatability of categorizing inherent translucency either in natural teeth or crowns exists.

The intention of this study was not set in having evaluators try and match translucency values generated by either an instrumental measurement or consensus of a formulating group of clinicians, but to evaluate the repeatability between dental faculty actively practicing restorative dentistry and final year dental students, categorizing inherent translucency of images selected at random using either a 3-point or 7-point scale. The null hypothesis is that categorizing inherent translucency of an anterior tooth by dental faculty whose clinical practice was primarily restorative dentistry and final year dental students using a 3-point scale versus a 7-point scale will provide comparable results.

2. Results

Scatter plot diagrams of the 3-point and 7-point scale results are presented in Figure 1, 3-point scale (1A faculty, 1B students) and Figure 2, 7-point scale (2A faculty, 2B students). The x axis showing the numbered slides (from 1 to 25) and the y axis showing the scale used (3 versus 7), from very translucent (1) to opaque (3 or 7 depending on scale).

In determining the inter-rater reliability of categorizing inherent translucency of anterior dentition, the closer the covariance parameter estimate (CPE) value is to zero, the more consistent the repeated selection. The overall combined CPE of the final year dental student and dental faculty who are actively practicing restorative dentistry can be seen in Table 1. The CPE for the 3-point scale was 0.054, while to CPE for the 7-point scale was 0.513. The parametric Pearson correlation coefficient 0.37 and non-parametric Spearman correlation coefficient 0.57, both elicited a positive relationship between the 3-point scale group and the 7-point scale group, using the 1200 data points available. To normalize the comparison between the 3-point and the 7-point scale, the Z value is used which is the CPE/Standard Error. The Pr > Z is the $p$ value, for the probability that the CPE is significantly equal to zero, other than random variables.
Table 1. Combined faculty and dental student covariance parameter estimates (CPE) for both scales.

<table>
<thead>
<tr>
<th>Faculty and Dental Students</th>
<th>CPE</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr &gt; Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-point scale</td>
<td>0.054</td>
<td>0.023</td>
<td>2.32</td>
<td>0.0101</td>
</tr>
<tr>
<td>7-point scale</td>
<td>0.513</td>
<td>0.129</td>
<td>3.98</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Figure 1. Scatter plot data for 3-point scale, (A)—faculty, (B)—students.
In determining the inter-rater reliability of categorizing inherent translucency of anterior dentition, the closer the covariance parameter estimate (CPE) value is to zero, the more consistent the repeated selection. The overall combined CPE of the final year dental student and dental faculty who are actively practicing restorative dentistry can be seen in Table 1. The CPE for the 3-point scale was 0.054, while the CPE for the 7-point scale was 0.513. The parametric Pearson correlation coefficient 0.37 and non-parametric Spearman correlation coefficient 0.57, both elicited a positive relationship between the 3-point scale group and the 7-point scale group, using the 1200 data points available. To normalize the comparison between the 3-point and the 7-point scale, the Z value is used which is the CPE/Standard Error. The Pr > Z is the p value, for the probability that the CPE is significantly equal to zero, other than random variables.

Table 1. Combined faculty and dental student covariance parameter estimates (CPE) for both scales.

<table>
<thead>
<tr>
<th>Group</th>
<th>CPE</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr &gt; Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty</td>
<td>1.80</td>
<td>0.114</td>
<td>15.74</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dental Student</td>
<td>0.16</td>
<td>0.019</td>
<td>7.98</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Five digital images were repeated twice in the study, the results of comparing the repeatability of reviewing the same slide more than once in the slide show using the 3-point scale can be seen in Table 2. Final year dental students were shown to be more consistent (CPE of 0.16) than dental faculty (CPE of 1.80) when evaluating the five repeat slides (p < 0.05). When reviewing the repeat slides more than once, using the 7-point scale, the results (as determined by the CPE) were identical as seen in Table 3, (both students and faculty had a value of 0.669).

Table 2. Covariance parameter estimates (CPE) for 3-point scale.

<table>
<thead>
<tr>
<th>Group</th>
<th>CPE</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr &gt; Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty</td>
<td>1.80</td>
<td>0.114</td>
<td>15.74</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dental Student</td>
<td>0.16</td>
<td>0.019</td>
<td>7.98</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Figure 2. Scatter plot data for 7-point scale, (A)—faculty, (B)—students.
When viewing the interaction between evaluators (Table 4), there was not a significant difference from zero (CPE = 0.021 ($p = 0.083$)). This shows that the evaluators were relatively consistent when making decisions on the inherent translucency for each slide. The CPE was significantly different from zero for Slides = 0.185 ($p = 0.001$) which is expected. There is an interaction between Evaluators x Slide interaction (CPE= 0.122 ($p = 0.0002$), which shows that different evaluators rated the inherent translucency of the image differently that the next evaluator.

### Table 4. Covariance parameter estimate (CPE) for subject and slides.

<table>
<thead>
<tr>
<th>Factors</th>
<th>CPE</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr &gt; Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluator</td>
<td>0.021</td>
<td>0.015</td>
<td>1.39</td>
<td>0.083</td>
</tr>
<tr>
<td>Slides</td>
<td>0.185</td>
<td>0.062</td>
<td>3.01</td>
<td>0.001</td>
</tr>
<tr>
<td>Evaluator * Slide</td>
<td>0.122</td>
<td>0.035</td>
<td>3.51</td>
<td>0.0002</td>
</tr>
<tr>
<td>Residual</td>
<td>0.180</td>
<td>0.031</td>
<td>5.77</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

—Interaction between Evaluators and Slides.

### 3. Discussion

The null hypothesis, in comparing the ability to categorize inherent translucency by dental faculty with an emphasis on restorative dentistry and final year dental students using the 3-point scale versus a 7-point scale provided a positive relationship, Spearman ($r = 0.57$) tests show moderate positive correlation coefficients. With a moderate positive relationship, the 3-point and 7-point scales could be used interchangeably but with caution. The problem in comparing the scales is that the 7-point scale had more categories available and cannot be compared positively with the 3-point scale with less variation. A 5-point scale was not used but that might be a compromise between the two scales that were selected in the current analysis.

The perception of slight changes in the inherent translucency of a single tooth can depend on the structure of the viewing scene [12]. Takasaki [13] and Whittle [14] both discuss the concept of the “crispening effect”. When objects are presented on uniform backgrounds, as the single tooth and consistent background used in the current survey, there is an increased apparent contrast between two colors of similar lightness against the lightness of value between them [15]. Natural occurring scenes are more complex with the background contrast being more varied and potentially different stimulus responses expected [12,16,17].

Lindsey and Wee [18] showed that when an individual is tasked with selecting a category when judging a slight color change, they are potentially biased to selecting “yes”, that a difference does exist, even when the object may not show any perceivable difference. There may be a false alarm rate in all selection studies [18]. One would expect the false alarm rates to be higher in natural cases because of outside factors that may influence one’s selection of a category.

Table 3 showed significantly more variability between evaluators than within evaluators. Although the CPE for the 3-point scale looks to be one-tenth of the 7-point scale, these scales cannot be readily compared due to the variability of the number of choices available to participants. There is only a small amount of variability attributed to the evaluator to evaluator variability that is not statistically significantly different (0.21 ($p = 0.083$)) (Table 4). Inter- and intra-rater reliability might be affected by the discrimination fineness of the data the evaluators must consider. To help the evaluators in the study, a
more in-depth description, on what constitutes variations in inherent translucency may need to be presented to students and faculty prior to participating in the study.

In this investigation, final year dental students were found to be more consistent than faculty when looking at repeat digital images using the 3-point scale, where on the 7-point scale they matched. The fact that the faculty was considerably less consistent with dental students on the 3-point scale versus matching on the 7-point scale may have been an aberration in this study. Inherent translucency is not a concept readily discussed and new clinical providers, like dental students, might be more accepting of new ideas and sensitive to evaluating using the methods prescribed. Faculty at a Creighton Dental School vary in age from 30s to 70s and come with variable backgrounds and experiences which might render too much confidence in the selection process.

No previous studies have tried to evaluate the repeatability of categorizing inherent translucency on some scale or chart. The consensus is to state that differences in inherent translucency occur naturally and should be considered a component of characterizing newly fabricated restorations. The images in the present study were two-dimensional, due to the fact that they were presented on a computer screen. Three-dimensional models may be more realistic and may have presented a different selection rate. Enlarging images on the monitor screen may also have an effect on the perceived inherent translucency of teeth.

Dental faculty and final year dental students were selected as evaluators for this study since they present a component of those individuals that would generally be selecting shades and translucency to present to the laboratory for fabrication of crowns. No laboratory technicians took part in the selection process, yet they are a very important component of the fabrication process of new restorations. Dentists can select a value for inherent translucency, but the laboratory technician must be able to distinguish between different values on a scale and fabricate a restoration to match the requested value. Although values were selected on either the 3-point or 7-point scale by dentists there are no models or pictures to designate the differences between the values for potential laboratory technicians. Additional scrutiny could have been accomplished to make sure digital slides matched each category on the 3- and 7-point scales. Use of a colorimetric system to actually measure and record values may have been an adjunct in setting up the scales. In this investigation digital photographs of clinical cases were used to try and get a sampling of reasonable values in a clinical setting. Small distractors were a limitation to this study. More idealization into selecting only restoration free teeth, abutting teeth without any restorations and a photographic environment without any opposing teeth or other structures in the background should be built into additional studies. The ideal scenario would be to accomplish the study in a clinical setting, using natural teeth, void of any distractors and using natural lighting versus digital images enlarged on a monitor screen.

The use of a 3-point scale may limit the selection range for a clinician or laboratory technician and the 7-point scale may give too many options. Possibly a 5-point scale would have been more ideal, once again, standards must be set on what the actual points on the scale represent.

Every evaluator used their own personal interpretation to construct the different categories on the 3- or 7-point scales. Ideally, there should be consistent values from which the laboratory technician will work. Fabrication of different three-dimensional models, similar to a shade guide, may make it easier for a technician to replicate a request from a dentist in the final restoration. Multiple indices have been used to calibrate translucency: total or direct transmission coefficient (tc), translucency parameter (TP) and contrast ratio (CR) [11]. In order to validate and fabricate ideal models, it may be ideal to construct tabs that meet either TP or CR values. Ideally, technicians should be part of the overall process in selecting a value on whichever system is used, since they are critical to restorative success.
4. Materials and Methods

The study was conducted in accordance with the Declaration of Helsinki, and the study protocol was approved by the Creighton University Institutional Review Board (#12-16493). A total of 40 human participants over the age of 18 years were recruited at the Creighton University School of Dentistry. A recruitment letter was sent by the study coordinator to potential participants via email. Interested dental students and faculty contacted the research laboratory and were screened to ensure they satisfied the inclusion and exclusion criteria for this study. The inclusion criteria included: twenty subjects, with equal gender balance, recruited from two categories: (1) final year dental students in clinical practice, and (2) dental faculty whose clinical practice was primarily restorative dentistry [18]. A power analysis was not carried out to determine the number of evaluators for this study. Since this is a pilot study, the actual delta was difficult to estimate, the number selected was arbitrary. Exclusion criteria included non-English speaking participants and individuals with colorblindness.

Each participant was provided a letter of participation that explained the purpose of the research study, the role of the participant, potential risks and benefits, measures to protect privacy, and the right to withdraw participation at any point during the study. The Bill of Rights for Research Participants (http://www.creighton.edu/fileadmin/user/ResearchCompliance/IRB/Policies and Procedures/118_11_Bill_of_Rights_for_Research_Participants.pdf) (accessed on 11 December 2012) was listed at the end of the letter. All subjects gave their informed consent for inclusion before they participated in the study. Participants were screened for color vision deficiency using the American Optical Company Hardy-Rand-Ritter (AO-HRR) pseudoisochromatic plate test [19].

Digital images of 25 single anterior teeth were selected from previous images of clinical patients taken under normal clinical situations (i.e., in a clinical operatory under natural lighting, no additional operatory light was projected on teeth at photographic appointment). Differences in age, gender or restorative treatment were not recorded. A digital clinical camera, Canon EOS 10D with macro ring lite MR-14EX flash was used to generate all digital photographs. The images were visually selected based on a range of potential inherent translucency by a single individual (AGW). The selection process was subjective, random, and meant to include examples of different values along the inherent translucency axis. No instrumental measurements were carried out to determine actual inherent translucency values nor did authors form a consensus on where they would occur on either the 3-point or 7-point scales. The objective of the study was not to have individuals validate inherent translucency, by matching a value on a colorometric scale, but to categorize the repeatability between evaluators. Images were enlarged and cropped so that primarily only one complete tooth with a slight amount of the adjacent dentition was visible (Figure 3). The images were constructed with a dark gray background using Microsoft Office Power Point (Microsoft Inc., Redmond, WA, USA). Either the 3- or 7-point scale was placed at the bottom of each image for reference. Five of the 25 dental images were repeated twice resulting in 30 digital images that were randomized for each participant. Digital projections were labeled (#1 to #30), were placed in the same order for a specific participant but randomized between participants. Figure 4 shows an example of an opaque tooth while Figure 5 shows an example of a very translucent tooth.
After determining participant (to be called evaluator) eligibility, dental faculty (n = 20) and final year dental students (n = 20) were randomized and placed into two groups. The evaluator was instructed to judge the overall translucency of the tooth image. The first group categorized the tooth images using the 3-point scale (very translucent, moderately translucent, or opaque) and two weeks later, the 7-point scale (1–7 with 1 as very translucent and 7 as opaque). Figure 1 shows an example representative slide presented to the viewing evaluator with the 3-point scale. The second group categorized the images using the 7-point scale first and two weeks later, the 3-point scale. The study design and progression are illustrated in Figure 6. Evaluators going through for the second time with the alternate point system did not have access to their previous results.
After determining participant (to be called evaluator) eligibility, dental faculty (n = 20) and final year dental students (n = 20) were randomized and placed into two groups. The evaluator was instructed to judge the overall translucency of the tooth image. The first group categorized the tooth images using the 3-point scale (very translucent, moderately translucent, or opaque) and two weeks later, the 7-point scale (1–7 with 1 as very translucent and 7 as opaque). Figure 1 shows an example representative slide presented to the viewing evaluator with the 3-point scale. The second group categorized the images using the 7-point scale first and two weeks later, the 3-point scale. The study design and progression are illustrated in Figure 6. Evaluators going through for the second time with the alternate point system did not have access to their previous results. The evaluation was performed in a dark room, without any external lighting, illuminated only by the light of the study monitor, a Sony Telemaster EL Professional Video Monitor (Sony Corp., Tokyo, Japan). The settings of the computer monitor were not altered during the duration of the study. The clinical coordinator had a separate computer monitor to control the images viewed by the participant. Evaluators were seated with their lower jaw and forehead resting lightly on a head frame (similar to what is used during an optometry examination) (Heavy Duty Chin Rest, Richmond Products, Albuquerque, NM, USA) that was attached to a Table 20.5 inches away from the screen (Figure 7).

Evaluators adapted to the dark gray background of the monitor for two minutes before data collection began. Evaluators verbally categorized 30 digital images of vital anterior teeth using either the 3-point scale or the 7-point scale, dependent on their assigned group. The clinical coordinator was present in the room guiding each evaluator through the presentation. Evaluator answers were recorded by the research assistant on the data collection sheet. There was no time constraint in place and the evaluator instructed the research assistant to move on to the next slide when ready. After a minimum of two weeks,
evaluators were reminded via email to return for the second session to categorize the images using the opposite scale.

![Subject positioning and setup for evaluating monitor image.](image)

Figure 7. Subject positioning and setup for evaluating monitor image.

Statistical analysis was performed using SAS statistical software (SAS Institute Inc. Cary, NC, USA). An F test for analysis of variance at 95% confidence interval was used to compare the variability and repeatability of the 3- and 7-point scales for the dental faculty and dental student groups. Pearson and Spearman correlation coefficients were used to measure the relationship between the 3- and 7-point scales. A covariance parameter estimate was accomplished to compare the inter-rater variability of the two variables, the dental faculty versus final year dental students.

5. Conclusions

Either a 3-point or 7-point scale can be used to categorize inherent translucency in teeth, although cannot be used interchangeably. When viewing the same slide more than once, final year dental students were more consistent than faculty using the 3-point scale but in accord with faculty using the 7-point scale.


Funding: Supported in part by a Health Future Foundation Development Grant, #24006, Creighton University.

Institutional Review Board Statement: The study was conducted according to guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Creighton University #12-16493 on 12/11/2012.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.
References