The Correlation of Student Performance in Preclinical and Clinical Prosthodontic Assessments

Donald A. Curtis, D.M.D.; Samuel L. Lind, Ph.D.; Sheila Brear, B.D.S.; Frederick C. Finzen, D.D.S.

Abstract: Tracking student performance in preclinical and clinical courses can be helpful in developing and refining a curriculum. Our objective was to correlate student performance on three fixed prosthodontic examinations taken by eighty junior dental students. Examinations included a knowledge-based objective structured clinical examination (OSCE), a manual skills exercise completed on a typodont (Typodont), and a competency casting exam (Casting CE) on a patient. Multiple regression analysis indicated that the OSCE and Typodont exam scores, as independent variables, were not statistically significant predictors (P=0.07; P=0.87, respectively) of Cast CE exam performance, which was the dependent variable. Correlations were weak for the OSCE (r=0.21) and nearly nonexistent for the Typodont exam(r=0.03) when compared to the Casting CE. Our results indicate a weak correlation between an OSCE-based knowledge exam measuring students’ knowledge of critical errors in preparations and castings and a competency exam involving the preparation of a full veneer crown. Results also indicate virtually no correlation between a typodont preparation examination designed to provide a measure of students’ clinical skill and a clinical competency exam involving the preparation of a full crown.

Dr. Curtis is Professor, Department of Preventive and Restorative Dental Sciences, School of Dentistry, University of California, San Francisco; Dr. Lind is Associate Professor, School of Economics and Business Administration, Saint Mary’s College of California; Dr. Brear is Health Sciences Assistant Clinical Professor and Division Chair, School of Dentistry, University of California, San Francisco; and Dr. Finzen is Health Sciences Clinical Professor and Chair of Prosthodontics, School of Dentistry, University of California, San Francisco. Direct correspondence and requests for reprints to Dr. Donald A. Curtis, Department of Preventive and Restorative Dental Sciences, D-3212, School of Dentistry, University of California, San Francisco, 707 Parnassus, San Francisco, CA 94143-0758; 415-476-5827 phone; 415-476-0858 fax; Don.Curtis@ucsf.edu.

Key words: prosthodontics, objective structured clinical exam, student assessment, student performance

Submitted for publication 8/14/06; accepted 10/30/06

A dental curriculum is a four-year opportunity for dental educators to progressively develop a student’s knowledge, skills, and attitudes with the hope of graduating a competent beginning dentist. As students progress through the dental curriculum, knowledge is incrementally gained from classroom, seminar, clinical, and peer interactions. Manual skills development occurs initially with simple preclinical bench-top typodont procedures and progresses to mannequin exercises before students enter clinics. There are threshold areas in a dental curriculum where assessments are completed to determine if a student has integrated knowledge and/or skills sufficiently before transitioning to the next level of the curriculum. An example of a threshold area in the fixed prosthodontics curriculum at the University of California, San Francisco (UCSF) is the examination to determine if a student has sufficient didactic knowledge and understanding of clinical procedures in fixed prosthodontics to transition into the clinics. A second example is for students to favorably complete a fixed partial denture (FPD) preparation on a typodont before transitioning further into clinical care. The purpose of this study was to determine the correlation between two threshold examinations in preclinical fixed prosthodontics and a clinical competency exam in fixed prosthodontics.

Predicting future professional achievement has been an elusive goal in dental education. Most measures used to predict achievement in preclinical and clinical activities have shown weak predictive values.1-12 Research has been completed evaluating the correlation of preadmission measures with preclinical performance,2,4,5,7,11,12 the correlation of preadmission measures with general clinical performance,9 and the correlation of assessments within or to a specific discipline.9,10
Of the research that has evaluated the value of preadmissions measures to predict student performance in preclinical courses, most have shown that the Perceptual Abilities Test (PAT) is a better predictor of preclinical technique course grades than preadmission college grade point average (GPA) or science grade point average (SGPA).\textsuperscript{4,7} However, the PAT remains only modestly correlated to student performance in preclinical courses.\textsuperscript{4,7}

While investigators have determined modest correlations between various preadmissions assessments and preclinical performance, only weak correlations have been shown between a student’s preadmissions background and overall clinical performance or performance in a specific dental specialty.\textsuperscript{9,11} Potter, in a study of two graduating dental school classes et al., showed that clinical performance correlated weakly with preadmission GPA, preclinical basic science GPA, or dental school lecture or laboratory performance.\textsuperscript{9} Bellanti et al., in a study of 344 students, correlated Dental Admission Test (DAT) scores to performance in fixed prosthodontics and showed that the DAT carving test and preadmissions GPA were better predictors of success in preclinical prosthodontics than other DAT subscores but were weak predictors of clinical prosthetic achievement.\textsuperscript{11} As part of a larger study comparing problem-based learning pedagogy to traditional teaching, Rich et al. reported that the variance of the score on a clinical periodontal assessment was minimally explained by the preclinical midterm and final exams.\textsuperscript{10} Rich et al. found the inability of the preclinical assessments to predict clinical performance in the same specialty was perplexing.

The objective structured clinical examination (OSCE) is a standardized approach to evaluate performance in simulated situations.\textsuperscript{13-15} In many curriculums the OSCE is used as a formative assessment to provide feedback to students.\textsuperscript{16} Some educators have argued that there is sufficient validity to the OSCE that it can be used as a “gateway event,” where failing students must complete remedial education before progressing.\textsuperscript{17} Unfortunately, few investigators have examined the OSCE’s predictive validity.\textsuperscript{18}

Although previous investigations have shown a modest predictive ability of preadmission assessments to professional achievement in dental school, few studies have evaluated the predictive value of preclinical performance in a specialty to clinical performance in the same specialty.\textsuperscript{10,11} Accordingly, the purpose of this investigation was to correlate two preclinical prosthodontic assessments (an OSCE assessment and a manual skills typodont preparation) with a clinical assessment of student performance in fixed prosthodontics.

**Methods**

Eighty-two students from the UCSF graduating class of 2005 were proposed for inclusion in this investigation. Two students did not participate in the study—one because of an extended illness and the other because of academic probation. Therefore, eighty students participated in this study. As part of their normal requirements, all students completed a knowledge-based objective structured clinical exam (OSCE), a skills-based typodont examination (Typodont), which included the preparation of a three-unit fixed partial denture (FPD), and a patient-based competency exam (Casting CE) that included a full veneer preparation, fabrication of a provisional, and delivery of a full crown.

Test results were collected with code numbers replacing names so confidentiality was maintained. The protocol for the study was reviewed and approved by the Committee of Human Research (CHR) at the University of California, San Francisco.

**Objective Structured Clinical Exam (OSCE)**

The OSCE typodont-based assessment included ten stations with twenty-seven total questions rendering sixty-three possible points. The OSCE assessment included two parts: five stations with questions about completed typodont FPD preparations and five stations with questions about competed FPD castings. At each station, students were provided mylar strips, explorer, mirror, floss, and articulating paper. Typodonts were placed in mannequins oriented in a dental chair to a reclined position that allowed direct vision, but students benefited from using a mirror for indirect vision.

Five stations of the OSCE included basic questions about FPD preparations completed on typodont teeth #18 and #20 with #19 as a pontic. Questions about taper of individual typodont teeth, draw, occlusal reduction, and margin placement were queried. Questions also included items related to an acrylic provisional restoration that was fabricated and included questions about why a provisional restoration may have fractured (narrow connector) or what was wrong with a provisional restoration (occlusion too high). Questions were formatted as multiple choice
questions having four possible and one correct answer to be identified.

The second part of the OSCE assessment included five stations where three unit cast FPDs had been fabricated and placed on prepared typodont teeth #18 and #20 with #19 as a pontic. Two or three critical errors were waxed and cast in each of the five FPDs that included the following: occlusion that lacked contact to the opposing arch, a casting with a facial margin that was short of the prepared margin, a contact that was binding and kept the FPD from seating, an open interproximal contact, damage of an adjacent tooth, and improper finishing of the restoration (Figure 1). The questions required a written response, and students were asked to evaluate each FPD problem and state if the casting was correctable or not. If the casting error was thought to be correctable, such as a missing mesial contact that could be soldered, the corrective action needed to be identified. One of the five cast FPDs was an ideal casting that students were asked to identify. The three unit castings were cast from type 4 gold (Jelenko, San Diego, CA).

The completed OSCE examinations were collected and coded by a staff member so that the faculty member grading the examination was blinded to the student’s identity. One faculty member graded all OSCE exams.

Manual Skills Examination (Typodont)

The manual skills examination included students’ preparing a three-unit posterior fixed partial denture (FPD) on typodont teeth followed by the fabrication of a provisional restoration. The FPD included full veneer crowns on #18 and #20 with #19 as a pontic. Typodont tooth #17 was present. Typodonts were placed in a mannequin and positioned in a reclined clinic chair to simulate patient care as closely as possible (Figure 2). The examination protocol also included following standard infection control procedures. Students were given three hours to complete the FPD
preparations and provisional restoration. Scores were given for eight subdivisions on the preparation and five on the provisional for a total possible score of 130 points.

The typodonts were removed by staff and graded by one of three calibrated faculty. Each exam was coded so the faculty grading the exam was blinded to the student’s identity. The completed preparations were evaluated for occlusal and axial reduction, margin placement, taper of individual preparations and presence of undercuts between preparations, assessment of adjacent teeth, and finish. The provisional restoration was evaluated for marginal adaptation, occlusion, proximal contacts, pontic contour, embrasures, and finish.

Casting Competency Examination (Casting CE)

The casting competency assessment involved assessing a student’s skills in preparing a tooth for a full crown, fabricating and adjusting a provisional, and delivering a permanent crown (Figure 3). The completed preparation and provisional restoration were assessed by calibrated faculty. Proper occlusal and axial reduction, retention and resistance form, appropriate taper and draw, margins, and assessment of adjacent teeth were completed. Additionally, the provisional was evaluated for proper contacts, occlusion, and finish. The score was a composite of ten sections of the exam with 100 points possible. The criteria for evaluating the casting competency examination were nearly identical to the criteria for evaluating the manual skills typodont examination.

Data Analysis

Descriptive statistical measures of the mean, standard deviation, range of scores, and coefficient of variation were determined for describing the characteristics of the OSCE assessment, the manual skills (Typodont) assessment, and the casting competency examination (Casting CE) (Table 1). The Coefficient of Variation (CV) provided a basis for a comparison

<table>
<thead>
<tr>
<th>Table 1. Descriptive statistical measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>Clinical Casting Exam</td>
</tr>
<tr>
<td>OSCE Exam</td>
</tr>
<tr>
<td>Typodont Exam</td>
</tr>
</tbody>
</table>
of the relative variation in the data. Scatter diagrams were plotted to visualize the distribution among the three variables (Figures 4 and 5).

Multiple regression analyses were completed with the OSCE and Typodont exams as independent variables and the Casting CE as the dependent variable. We tested several hypotheses using exhaustive iterations of the independent variables as predictors of the dependent variable, including 1) OSCE as a single independent variable, 2) Typodont as a single independent variable, and 3) OSCE and Typodont as multiple independent variables. Multiple correlations were calculated among the three data sets using Pearson’s product-moment correlation coefficients (Tables 2 and 3).

The overall Kappa statistic to quantify the level of inter-rater reliability was calculated.

**Results**

For the OSCE examination, the mean score was 38 ±7.2 with a range of 19 to 55. For the Typodont preparation examination, the mean score was 82
Table 2. Correlation analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Clinical Casting Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Casting Exam</td>
<td>1.00</td>
</tr>
<tr>
<td>OSCE Exam</td>
<td>0.21</td>
</tr>
<tr>
<td>Typodont Exam</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 3. Multiple regression analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSCE Exam</td>
<td>0.07</td>
</tr>
<tr>
<td>Typodont Exam</td>
<td>0.87</td>
</tr>
</tbody>
</table>

±24.8 with a range of 22 to 128. For the casting competency clinical examination, the mean score was 92 ±6.1 with a range of 71 to 100. The Coefficient of Variation, a comparison measure of relative dispersion, showed that the Casting CE series varied in a narrow range (7 percent) while the OSCE (19 percent) and Typodont (30 percent) data were much more volatile. Scatter plots provided a visual representation of these differences in the data variation (Table 1, Figures 4 and 5).

Multiple regression analysis indicated that the independent variables (OSCE P=0.07; Typodont P=0.87) were not statistically significant predictors for the dependent variable, Casting CE (Table 3). Alternative regression models of the independent variables also demonstrated that, individually, neither was a significant predictor of the dependent variable. Correlation coefficients between Casting CE and OSCE (r=0.21), between Casting CE and Typodont (r=0.04), and between OSCE and Typodont (r=0.08) further confirmed virtually no relationship among the variables (Table 2). Notwithstanding the lack of statistical significance for the independent variables, coefficient of determination values (R-Squared: 4.5% OSCE, 0.1% Typodont) suggested no explanatory power of the independent variables for the dependent variable. The lack of association among the variables can be observed on scatter diagrams of the variables (Figures 4 and 5).

The overall Kappa statistic to quantify the level of inter-rater reliability was completed for the Typodont exam (0.78). One faculty member graded all OCSE exams so a Kappa score was not calculated for the OSCE exam.

Discussion

Our results indicate a weak at best correlation between a preclinical OSCE-based exam measuring a student’s ability to identify critical errors in preparations and castings and a clinical competency exam involving the preparation for and delivery of a full veneer crown. This finding was surprising to us. We would have expected that students doing well on the OSCE exam, with demonstrated skills to detect errors in castings and preparations, would be better prepared to complete a more favorable result on the casting competency examination. However, our findings are consistent with Potter et al. who found no significant correlation between preclinical grades and clinical grades. Additionally, our findings are consistent with Rich et al., who demonstrated a low correlation between preclinical assessments of student performance and clinical performance in the same specialty (periodontics). Although some educators have argued that the OSCE has sufficient validity to be used as a threshold examination, the results of our study would lead us to not consider using the OSCE as a “must pass” before continuing in the curriculum. Our recommendations for the use of the OSCE in our curriculum would be consistent with Mavis et al., who felt the OSCE is best utilized as a formative assessment by which to provide feedback to students rather than as a threshold examination.

Our results also indicate a nearly nonexistent correlation between a preclinical typodont preparation examination designed to provide a measure of a student’s psychomotor skill and a clinical competency exam involving the preparation of a full crown. This also surprised us. We expected that students who did well on a psychomotor-based skills examination in a preclinical exercise would be better prepared to do well on a clinical casting competency examination, especially examinations within the same specialty. The reasons for a low correlation between student performance in preclinical laboratory performance and in clinical assessments have been explored by Chambers, who suggested that the preclinical and clinical contexts differ substantially and factors that contribute to success in each may also differ. Chambers also suggested that overlearning, or practicing a skill beyond the point of initial correct performance, may be helpful to transfer skills from preclinical to clinical contexts.
The question has to be asked why the preclinical exams are not more predictive of future clinical performance. Chambers addressed this question by suggesting that the preclinical and clinical contexts differ and that skills such as communication, decision making, and management are new issues that students entering the clinics must face and are not issues in the consistent and tightly controlled preclinical environment. This concept is supported by Zakay and Wooler, who found that training conducted under normal conditions improved decision performance under normal conditions but did not improve performance under conditions of reduced time. It could be that the preclinical and clinical conditions are sufficiently different, as Chambers has suggested, that the conditions for success require different skill sets.

One might expect a stronger correlation of the OCSE to the Casting CE if the format of our OSCE more closely simulated patient care. In a very interesting study of 137 medical students, Wilkinson and Frampton found that a clinically based OSCE was a stronger predictor of subsequent clinical performance than traditional essay and multiple choice questions. However, they also found that combining the traditional essay and multiple choice exam increased predictive validity, something our results did not show.

Although the preclinical exams are not significant predictors of clinical performance, the preclinical exercises and exams still have merit in providing repetition on tasks similar to the clinical experience. Chaiken et al. have shown that psychomotor abilities do improve with practice and that processing speed, or the strength in the ability to complete visual extrapolation tasks, and time estimation tasks were the most important predictors for improvement of psychomotor skills. This may explain why investigators have shown that the PAT is more predictive of preclinical student performance than many other measures.

The coefficient of variation (CV) allows a comparison of the variation in the data and showed the Casting CE varied in a very narrow range (7 percent) while the OSCE (19 percent) and Typodont (30 percent) data were much more variable. The wide dispersion of the preclinical assessments may be a result of students not placing as much emphasis on preclinical exams and being willing to accept a lower grade. The narrow CV with the Casting CE could be related to faculty reluctance to give a low grade on a high stakes assessment, as well as the possibility the students had become proficient by the time they took the Casting CE.

An important assumption of regression analysis of multicollinearity is that there should be no highly correlated relationship between two or more of the predictor variables. Multicollinearity makes it difficult to obtain unique estimates of the regression coefficients because there are an infinite number of coefficients that would satisfy the regression model. Multicollinearity was tested for and ruled out by the multiple correlation analysis \( r = 0.08 \) and by modeling the two independent variables separately. As such, no multicollinearity was observed.

Several variables may have confounded our results. The same faculty did not grade all the examinations although faculty participated in calibration activities. The order of testing can have an influence on results, which might have occurred because the OSCE exam was given to small groups over a month-long period. It is possible that test information may have been shared by students who had completed the OSCE exam.

Future work will include continual evaluation of student assessments in hopes of providing a balance of formative and summative criteria by which objective and valid feedback to students can be provided.

REFERENCES