Agreement Among Dental Students, Peer Assessors, and Tutor in Assessing Students’ Competence in Preclinical Skills

Jennifer I. Foley, PhD; Gillian L. Richardson, MClinDent; Joyce Drummie, Dip Dent Therapy

Abstract: The aim of this study was to determine the level of agreement regarding assessments of competence among dental students, their student peers, and their clinical skills tutors in a preclinical skills program. In 2012-13 at the University of Edinburgh, second-year dental students learned to perform the following seven cavity preparations/restorations on primary and permanent Frasaco teeth: single-surface adhesive occlusal cavity; single-surface adhesive interproximal cavity; single-surface adhesive labial cavity; multi-surface adhesive cavity; multi-surface amalgam cavity; pre-formed metal crown preparation; and composite resin buildup of a fractured maxillary central incisor tooth. Each student, a randomly allocated student peer, and the clinical skills tutor used standardized descriptors to assign a competency grade to all the students’ preparations/restorations. The grades were analyzed by chi-square analysis. Data were available for all 59 second-year students in the program. The results showed that both the students and their peers overestimated the students’ competence compared to the tutor at the following levels: single-surface adhesive occlusal cavity ($\chi^2=10.63, p=0.005$); single-surface adhesive interproximal cavity ($\chi^2=11.40, p=0.003$); single-surface labial cavity ($\chi^2=23.70, p=0.001$); multi-surface adhesive cavity ($\chi^2=12.56, p=0.002$); multi-surface amalgam cavity ($\chi^2=38.85, p=0.001$); pre-formed metal crown preparation ($\chi^2=40.41, p=0.001$); and composite resin buildup ($\chi^2=57.31, p=0.001$). As expected, the lowest levels of agreement occurred on the most complicated procedures. These findings support the need for additional ways to help students better self-assess their work.

Dr. Foley is Senior Clinical Lecturer/Honorary Consultant in Pediatric Dentistry, University of Edinburgh Dental Institute; Ms. Richardson is Consultant/Honorary Senior Lecturer in Pediatric Dentistry, University of Glasgow Dental Hospital; and Ms. Drummie is Clinical Teaching Fellow in Pediatric/Restorative Dentistry, University of Aberdeen Dental School and Hospital. Direct correspondence to Dr. J.I. Foley, University of Edinburgh, Edinburgh Dental Institute, Lauriston Building, Lauriston Place, Edinburgh, EH3 9HA, United Kingdom; +44 131 536 3971; jfoley10@exseed.ed.ac.uk.

Keywords: dental education, restorative dentistry, clinical skills, clinical competence, self-assessment

Submitted for publication 2/23/15; accepted 5/1/15

The University of Aberdeen Bachelor of Dental Surgery (BDS) program is based on recommendations of both the General Dental Council (GDC) in the United Kingdom and the Association for Dental Education in Europe (ADEE) regarding the content and competencies for the curriculum. The program combines didactic and clinical skills education over the duration of the degree. Students’ exposure to hands-on clinical dentistry begins in the first three months of the second-year curriculum with an introductory Clinical Skills Program. In this program, students are introduced to rotary instruments and asked to perform specific simple treatments on patient simulators. In addition to formative clinical skills tutor-led assessment, students are encouraged to reflect on their own performance to enhance the learning experience and its effect on their practical skills.

A number of studies have found that simulation clinics improve student performance in practical skills although the majority of these have investigated the single-surface occlusal cavity. A few others have investigated full gold crown and full metallic-ceramic crown preparations and found variable results when comparing the grades awarded by experienced and peer assessors. One explanation for such disparities is the differing methods of assessment for different techniques, with studies using varying descriptors for the assessments, thus making direct comparisons between studies problematic. In addition, students often have varying experiences of self-reflection: those with greater insight into their abilities tend to underrate their performance, while students who are less aware of their deficiencies tend to overrate their performance.

In relation to assessment of skills, the GDC’s Preparing for Practice emphasizes that one of the learning outcomes for the undergraduate dental curriculum is to acknowledge the “importance of assessment, feedback and critical reflection, identification of learning needs, and appraisal in personal development planning.” The GDC’s Standards for
Education also places an emphasis on reflection and receiving feedback in student assessment processes.\textsuperscript{15} In the United States, the Commission on Dental Accreditation (CODA) standards require that graduates be able to demonstrate the capacity to self-assess and self-reflect and to evaluate their own competence as well as have the ability to provide feedback to and critique of fellow students.\textsuperscript{16}

The ability to assess one’s own performance is vital for any clinical discipline although there is evidence from previous studies that students tend to overestimate their clinical abilities.\textsuperscript{17-20} Peer assessment has also been widely advocated, with one study on student competence in third molar surgery demonstrating similar assessment scores between peer assessors and more experienced clinical trainers.\textsuperscript{19} To date, there is a paucity of data regarding self-, peer, and tutor assessment in preclinical skills programs; therefore, the aim of this study was to determine agreement about competence amongst these individuals.

Materials and Methods

Ethics approval was obtained from the University of Aberdeen’s College Ethics Research Board (CERB no. 1012) prior to study commencement. The study was undertaken with all second-year dental students during three introductory clinical skills courses at the University of Aberdeen between September 2012 and December 2013.

In the sessions, practical skills on primary and permanent Frasaco teeth mounted in Frasaco mannequin heads (Frasaco, Greenville, NC, USA) were taught. The seven cavity preparations/restorations taught were as follows: single-surface adhesive occlusal cavity; single-surface adhesive interproximal cavity; single-surface adhesive labial cavity; multisurface adhesive cavity; multi-surface amalgam cavity; pre-formed metal crown preparation; and composite resin buildup of a fractured maxillary incisor tooth. Prior to each session, a demonstration of the technique was provided. In addition, a PowerPoint presentation outlining the key components of each topic was placed on the students’ virtual learning environment 48 hours before the session.

On average, students were scheduled to 16 sessions in the skills room and were asked to complete each task twice. Due to schedule complications, however, some students were not available on the days when tasks were undertaken. As such, not all students were able to complete the requisite tasks, which accounts for the discrepancy in preparations/restorations presented.

In the assessment process, each student graded his or her own preparations/restorations according to a structured evaluation rubric for each procedure, all of which had been piloted and refined in the previous year of the course. For each procedure, ratings descriptors were given, and each was graded as either competent or noncompetent. The number of ratings descriptors for each skill was as follows: single-surface adhesive occlusal cavity: 8; single-surface adhesive interproximal cavity: 6; single-surface adhesive labial cavity: 7; multi-surface adhesive cavity: 9; multi-surface amalgam cavity: 10; pre-formed metal crown: 4; and composite resin buildup: 6. To achieve overall competence, all descriptors for each procedure had to be included. An example of one of the assessment rubrics is shown in Table 1.

Using the same rubrics, students were randomly assigned to independently review another student’s preparation/restoration for each skills assessment.

| Table 1. Evaluation rubric for pre-formed metal crown preparation |
|------------------------|------------------------|------------------------|
| Evaluation Criteria     | Competent              | Not Competent          |
| Mesial interproximal slice | • 15-20° taper        | • Taper < or > 15-20°   |
|                        | • Contact point cleared| • Contact point not cleared |
|                        | • Chamfer margin       | • Step margin           |
| Distal interproximal slice | • 15-20° taper        | • Taper < or > 15-20°   |
|                        | • Contact point cleared| • Contact point not cleared |
|                        | • Chamfer margin       | • Step margin           |
| Occlusal preparation    | • Reduced by 1.5 mm    | • Reduced by < or > 1.5 mm |
|                        | • Follows occlusal contour | • Flattened occlusal surface |
| Finish                 | • Smooth preparation   | • Sharp angles          |
|                        |                        | • Rough finish          |
Finally, all work was assessed using the rubrics by one of three clinical skills tutors, who are listed on the General Dental Council Specialist Register in Pediatric Dentistry. One-fourth of the preparations/restorations were randomly reviewed by the same clinical skills tutor one month after the initial evaluation. Again, an overall competent grade was assigned for each, and Kappa statistics were used to determine the clinical skills tutor reliability (Minitab 16.1.0).

Results

All 59 second-year students in the program participated in the study (F: 36; M: 23). The number of cavity preparations they completed were as follows: single-surface adhesive occlusal: 116; single-surface adhesive interproximal: 82; single-surface adhesive labial: 118; multi-surface adhesive: 118; multi-surface amalgam: 118; pre-formed metal crown: 54; and composite resin buildup of a fractured incisor tooth: 78.

Regarding assessment of competence for individual procedures, both students and their peers overestimated their competence compared to the clinical skills tutor for all the cavity preparations/restorations. Individual results were as follows: single-surface adhesive occlusal cavity ($\chi^2=10.63$, $p=0.005$); single-surface adhesive interproximal cavity ($\chi^2=11.40$, $p=0.003$); single-surface labial cavity ($\chi^2=23.70$, $p=0.001$); multi-surface adhesive cavity ($\chi^2=12.56$, $p=0.002$); multi-surface amalgam cavity ($\chi^2=38.85$, $p=0.001$); pre-formed metal crown preparation ($\chi^2=40.41$, $p=0.001$); and composite resin buildup ($\chi^2=57.31$, $p=0.001$) (Table 2). The intra-examiner reliability Kappa scores for the tutor’s initial evaluation and the evaluation one month afterwards were

Table 2. Proportion of cavity preparations deemed competent by students, their peers, and their clinical skills tutor

<table>
<thead>
<tr>
<th>Cavity Preparation/Restoration</th>
<th>Students</th>
<th>Student Peers</th>
<th>Tutor</th>
<th>$\chi^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-surface adhesive occlusal cavity</td>
<td>88%</td>
<td>88%</td>
<td>73%</td>
<td>10.63</td>
<td>0.005</td>
</tr>
<tr>
<td>Single-surface adhesive interproximal cavity</td>
<td>78%</td>
<td>90%</td>
<td>71%</td>
<td>11.40</td>
<td>0.003</td>
</tr>
<tr>
<td>Single-surface adhesive labial cavity</td>
<td>76%</td>
<td>88%</td>
<td>58%</td>
<td>23.70</td>
<td>0.001</td>
</tr>
<tr>
<td>Multi-surface adhesive cavity</td>
<td>56%</td>
<td>76%</td>
<td>54%</td>
<td>12.56</td>
<td>0.002</td>
</tr>
<tr>
<td>Multi-surface amalgam cavity</td>
<td>76%</td>
<td>80%</td>
<td>42%</td>
<td>38.85</td>
<td>0.001</td>
</tr>
<tr>
<td>Pre-formed metal crown preparation</td>
<td>78%</td>
<td>89%</td>
<td>50%</td>
<td>40.41</td>
<td>0.001</td>
</tr>
<tr>
<td>Composite resin buildup/fractured maxillary incisor tooth</td>
<td>87%</td>
<td>87%</td>
<td>46%</td>
<td>57.31</td>
<td>0.001</td>
</tr>
</tbody>
</table>

$^a$Statistically significant difference between students and tutor
$^b$Statistically significant difference between student peers and tutor
$^c$Statistically significant difference between students and student peers

Discussion

Traditionally, student performance in preclinical skills programs was assessed by a faculty member or tutor with appropriate feedback as part of formative assessment. Of late, however, students have been encouraged to reflect on their own performance, and peer feedback is being increasingly used in health professions education. After peer and self-assessment was introduced into the University of Edinburgh’s preclinical skills course, this study was designed to determine the degree of variation between the students’ and a clinical skills tutor’s assessments of cavity preparations/restorations performed on primary and permanent Frasaco teeth.

The results showed discrepancies between the competency assessment of the students and their peers and that of the tutor. Similar differences between students and instructors have been reported in previous studies. Recently, one group of researchers demonstrated a comparable finding whereby students overrated their performance quality compared to the staff rating on cavity preparation in a clinical skills setting. Another study found that students overrated their performance in a wax carving course in dental anatomy compared to faculty members. Indeed, there is evidence from multiple disciplines that individuals tend to overestimate their performance and those least skilled show the least ability to self-assess.
Our study found there were discrepancies in agreement on the competency scores for nearly all the procedures. In comparing the students’ scores with those of the tutor, we found discrepancies on five of the seven procedures; the exceptions were the single-surface adhesive occlusal cavity preparation and the multiple-surface adhesive restoration. In comparing the scores given by the students’ assigned peers with those of the tutor, we found discrepancies on six of the seven procedures; the exception was the single-surface adhesive occlusal cavity preparation.

In particular, there was a large discrepancy in the level of agreement for the composite resin buildup of the fractured permanent incisor tooth between both the students/student peers and the tutor. It could be argued that this was a more difficult task than cavity preparations, given the need to use a range of techniques, including preparing a cellulose acetate matrix, incremental packing of the matrix, and restoration placement and finish compared with the use of rotary instrumentation only for cavity preparation. Another study found that the more difficult the task for a cavity preparation on a dental manikin, the greater the level of disagreement between students and their tutor, with students overrating their performance compared to the tutor.\(^\text{12}\)

In addition, the degree of discrepancy may be attributable to the differing quantity and quality of previous learning experiences of the tutor and the students. The tutors, for example, would be well aware of the criteria by which cavity preparations/restorations would be judged since they created the descriptors and used them in previous clinical skills programs, whereas students were being introduced to the criteria for the first time. Other studies have found that students were less critical in grading their performance when they had less expertise than their instructor.\(^\text{13,26}\) Future research in skills training could examine differences in competency scores at various time-points throughout a skills program, given the assumption that students will progress in their abilities from the beginning through completion.

Regarding the level of agreement between the students and their peers, the assigned peers overestimated the students’ competency scores for all adhesive preparations except the single-surface occlusal one. Interestingly, however, for the two tasks deemed the most technically difficult (the amalgam multi-surface cavity preparation and the composite resin buildup), the two groups gave similar scores. It could be argued that student peers wished to ingratiate themselves with their fellow students by grading favorably, although each student peer was randomly allocated to review another student’s work for each procedure and this randomization would have decreased the likelihood that personal relationships would affect scoring. Interestingly, in another study that researched the reliability of self and peer scores in third molar surgery compared with tutors’ score, no statistically significant difference was found between tutor and peer assessment although the self-assessment scores were greater than the peer assessment scores, which is opposite to the findings in our study.\(^\text{19}\) Another group of researchers found that the nature of the relationship between students and their peer assessors could have an effect, with their grades of each other varying according to friendships or competing interests.\(^\text{27}\) To remove the possibility of collusion in our study, peers were randomly allocated for the assessments, and descriptors on a scoring rubric were used to minimize discrepancies in scoring. Given the differences between our findings and those of these previous studies, the effect of interpersonal factors on grading appears to be an area needing further study.\(^\text{21}\)

One of the limitations of our study was possible evaluation bias since one tutor did all the scoring. Although three tutors participate in this course, only one was available to take part in the study by scoring the preparations/restorations, given the volume of work required in the assessment process and the prior commitments of the other tutors. To avoid evaluation bias, further work in this area should include the recruitment and calibration of two faculty members who are not involved as clinical skills tutors in this program. In addition, since this study took place at one dental school and with only one group of students, the findings may not be generalizable to others.

**Conclusion**

The ability to reflect on and assess their own work is an essential skill for dental students to develop. This study sought to determine the level of agreement regarding assessments of competence among dental students, their student peers, and their clinical skills tutors in a preclinical skills program. The results showed that both the students and their peers overestimated the students’ competence compared to the tutor on nearly all the restorations/preparations included in the study. These findings contribute to dental educators’ understanding of how students assess their own and their fellow students’
skills in a preclinical lab setting and support the need for interventions throughout dental school to help students improve their ability to self-assess their work.

REFERENCES