Pilot Study: Agreement Between Self-, Peer- and Tutor- Assessment in a Pre-Clinical Skills Programme

Lubna Al-Muzian (✉ drlibi@hotmail.com)

Glasgow Orthodontic Academy   https://orcid.org/0000-0003-3136-1231

Antoniella Busuttil-Naudi

The University of Edinburgh

Jennifer Foley

The University of Edinburgh

Abstract

PURPOSE/OBJECTIVES: As dental education evolves, a shift towards self- and peer-assessments, producing dental professionals who are reflective, life-long learners, is evident. This study aims to establish the agreement of competency between and amongst students, and their course tutors, during a pre-clinical skills programme.

METHODS: Following cavity preparation of plastic teeth during a pre-clinical skills session, students assessed their own cavity preparations and those of their peers. Two course tutors undertook assessment of the same prepared teeth.

Plastic teeth were used during the pre-clinical skills sessions and assessed using a WHO CPITN probe and light.

Overall scores of competent (C) or not competent (NC) was assigned to each tooth. The data were collected and analysed using Fischer's exact test.

RESULTS: No statistical significance was found between self-, peer- or tutor-assessments. Tutors' assessments were in agreement. Peer- and self-assessment agreed with tutor-assessment two-thirds of the time. When not in agreement with tutor-assessment, peer- and self-assessment had a tendency to underestimate scores.

CONCLUSION: Whilst no statistical significance was found, tutor-assessment had a high degree of agreement. Peer- and self-assessment appeared to agree with the tutor-assessment two-thirds of the time. When not in agreement with the tutor, students tended to underestimate themselves and their peers.

Background

Dental/medical education is a field which is continually evolving. Over the past few decades, the focus has shifted from summative assessment towards methods of formative assessment. This change will meet the needs of the dental/medical profession in the long-term to produce more reflective clinicians and, in turn, lifelong learners.[1–4]

The three main methods of assessment are self-, peer- and tutor-assessment. The ideal of producing lifelong, reflective learners appears simple in theory, though in practice the self-reflection necessary to drive an adult towards the learning needed to improve their skills may not be so straightforward.

Some studies have shown that most professionals are not good at assessing themselves; however, these reviews determined that the ability to self-assess is key in becoming, and remaining, a successful professional in the dental/medical field.[2,5–10] On the other hand, peer-assessment appears to hold more weight in terms of validity and is usually in line with tutor ratings.[5,6,11–17]

Tutor-assessment, both summative and formative, is the conventional method of assessment.

Pre-clinical skills laboratory course component forms an essential part of the Bachelor of Science in Oral Health Science (BSc OHS) degree at most universities within the United Kingdom. Currently, within Edinburgh, tutor-assessment of the pre-clinical session is utilised, although no self- or peer-assessments are included in this component.

The comparison between modes of assessment has been a topic of research over the past few decades. In the field of dentistry, specifically in the area of pre-clinical skills, there is limited research which compares the different assessment outcomes. A study in 2001, found that the methods of self- and peer-assessment are under-utilised in the clinical training of dentists in the restorative field.[18] With the drive for self-directed learning in full swing, it is important not only to impress these methods of assessment onto clinical programmes in dentistry, but also to ensure their proper execution and effective utilisation.

Another study compared the self-, peer- and tutor- assessments in relation to a clinical procedure and found, in line with most research, that self-assessment had poor correlation with tutor scores. Peer-assessment scores, however, were consistent with tutor scores. They utilised checklists and global scoring which appears to give more reliable results in most studies.[15]

In 2015 a study undertaken in Dundee compared the assessment outcomes of self-, peer- and tutor-assessment in a pre-clinical laboratory setting. This study found that students needed more training to enhance both their self- and peer-assessment skills, as there was poor correlation between the three assessment types, with students over-estimating their own, as well as their peers, performance in comparison to the tutor's evaluation.[19] This outcome would suggest that students need more training in assessment, or that they do not hold the level of expertise required to measure and evaluate to a satisfactory level.

This project aims to encourage the incorporation of self- and peer-assessment into the pre-clinical curriculum as an adjunct to the conventional tutor-assessment in the pre-clinical laboratory component of the BSc OHS course.

Methods

Ethical approval was obtained from The University of Edinburgh, CMVM EREC (College of Medicine and Veterinary Medicine Educational Research Ethics Committee).

Students were recruited by distributing information regarding the study a week prior to the data collection sessions, and written consent forms were completed on the day of the data collection. At the planning stage of the study there were ten students eligible to progress to second year and participate in this research project, however, at the time of data collection only six students remained on the course.

In 2017–2018, second-year BSc OHS students performed four cavity preparations/restorations on primary and permanent plastic teeth: a single-surface adhesive occlusal cavity preparation and

restoration; a single-surface adhesive labial cavity preparation and restoration; a multi-surface adhesive cavity preparation and restoration, and; a multi-surface amalgam cavity preparation and restoration.

All prepared teeth were placed in a bag labelled with the student's university assigned code, Undergraduate Graduate (UG) followed by a two-digit number and stored.

Once the ethical approval was obtained in November 2017, each individual tooth was further labelled with a code which consisted of four components: (1) UG number (e.g. UG67); (2) Dentition: 01 (primary dentition), 02 (permanent dentition); (3) Type of cavity: 01 (single-surface adhesive occlusal cavity preparation and restoration); 02 (single-surface adhesive labial cavity preparation and restoration); 03 (multi-surface adhesive cavity preparation and restoration); 04 (multi-surface amalgam cavity preparation and restoration) and; (4) Serial number of tooth prepared: 01, 02, 03, 04, etc.

An example of the code is: UG67010301, this would refer to a tooth prepared by student UG67 (UG67), primary dentition (01), multi-surface adhesive cavity preparation and restoration (03) and it was the first tooth prepared in that manner (01). The next tooth to be identified as a multi-surface adhesive cavity preparation and restoration in the primary dentition by student UG67 would be numbered as: UG67010302.

Data Collection was carried out in two phases: (1) Student phase (March 2018): using standardised descriptors the students assessed their own prepared plastic teeth and those of a randomly allocated student peer, (2) Tutor phase (June 2018): two clinical skills tutors also assessed the same teeth using the same descriptors.

Each student assessed their own prepared plastic teeth as a self-assessment. Using a computer randomisation technique on Excel, the student's university numbers were randomly cross matched to select the student whose teeth they would assess as peers.

The plastic teeth to be assessed were given to each student along with the data collection sheet. They were asked to assess the teeth presented to them according to the criteria on the data collection sheet using a WHO CIPTN probe to aid them in identifying the characteristics of the tooth to be measured (cavity depth/ width, etc) and a light in the clinical skills laboratory.

The second phase of data collection involved two of the BSc OHS tutors assessing the same teeth in the same manner and completing data collection sheets accordingly.

The data was analysed by Fischer's exact test (IBM SPSS Statistics version 22) and the p-value was set at 0.05.

Results

Five students and two tutors took part in this project.

No statistical significance was found in the data when comparing self- to tutor-assessment, or peer- to tutor assessment, or self- to peer-assessment (Table 1 – 8). Further analysis using the data from the second tutor was not undertaken as the tutor-assessments were in agreement. However, patterns could be seen from the raw data.

Table 1
Fischer's Exact Test Applied to Assessment Results of Primary Single-Surface Adhesive Occlusal Cavity Preparations/ Restorations

	Fischer's Exact Test Statistic	p-Value
Self vs. Peer	1.0	1.0
Self vs. Tutor 1	0.444	1.0
Peer vs. Tutor 1	0.183	1.0
Self vs. Tutor 2	0.471	1.0
Peer vs. Tutor 2	0.206	1.0
Tutor 1 vs. Tutor 2	1.0	1.0
Nil significant		

Table 2
Fischer's Exact Test Applied to Assessment Results of Primary Single-Surface Adhesive Labial Cavity Preparations/ Restorations

	Fischer's Exact Test Statistic	p-Value
Self vs. Peer	1.0	1.0
Self vs. Tutor 1	0.471	1.0
Peer vs. Tutor 1	1.0	1.0
Self vs. Tutor 2	1.0	1.0
Peer vs. Tutor 2	0.559	1.0
Tutor 1 vs. Tutor 2	1.0	1.0
Nil significant		

Table 3
Fischer's Exact Test Applied to Assessment Results of Primary
Multiple Surface Adhesive Cavity Preparations/ Restorations

	Fischer's Exact Test Statistic	p-Value
Self vs. Peer	1.0	1.0
Self vs. Tutor 1	0.080	1.0
Peer vs. Tutor 1	1.0	1.0
Self vs. Tutor 2	0.077	1.0
Peer vs. Tutor 2	0.282	1.0
Tutor 1 vs. Tutor 2	1.0	1.0
Nil significant		

Table 4
Fischer's Exact Test Applied to Assessment Results of Primary
Multiple Surface Amalgam Cavity Preparations/ Restorations

	Fischer's Exact Test Statistic	p-Value
Self vs. Peer	0.524	1.0
Self vs. Tutor 1	0.167	1.0
Peer vs. Tutor 1	1.0	1.0
Self vs. Tutor 2	0.167	1.0
Peer vs. Tutor 2	1.0	1.0
Tutor 1 vs. Tutor 2	1.0	1.0
Nil significant		

Table 5
Fischer's Exact Test Applied to Assessment Results of
Permanent Single-Surface Adhesive Occlusal Cavity
Preparations/ Restorations

	Fischer's Exact Test Statistic	p-Value
Self vs. Peer	0.637	1.0
Self vs. Tutor	1.0	1.0
Peer vs. Tutor	0.619	1.0
Self vs. Tutor 2	1.0	1.0
Peer vs. Tutor 2	0.619	1.0
Tutor 1 vs. Tutor 2	1.0	1.0
Nil significant		

Table 6
Fischer's Exact Test Applied to Assessment Results of
Permanent Single-Surface Adhesive Labial Cavity Preparations/
Restorations

	Fischer's Exact Test Statistic	p-Value
Self vs. Peer	0.358	1.0
Self vs. Tutor	0.637	1.0
Peer vs. Tutor	1.0	1.0
Self vs. Tutor 2	0.335	1.0
Peer vs. Tutor 2	1.0	1.0
Tutor 1 vs. Tutor 2	1.0	1.0
Nil significant		

Table 7
Fischer's Exact Test Applied to Assessment Results of
Permanent Multiple Surface Adhesive Cavity Preparations/
Restorations

	Fischer's Exact Test Statistic	*p*-Value
Self vs. Peer	1.0	1.0
Self vs. Tutor	1.0	1.0
Peer vs. Tutor	1.0	1.0
Self vs. Tutor 2	1.0	1.0
Peer vs. Tutor 2	1.0	1.0
Tutor 1 vs. Tutor 2	1.0	1.0
Nil significant		

Table 8
Fischer's Exact Test Applied to Assessment Results of
Permanent Multiple Surface Amalgam Cavity Preparations/
Restorations

	Fischer's Exact Test Statistic	*p*-Value
Self vs. Peer	1.0	1.0
Self vs. Tutor	0.356	1.0
Peer vs. Tutor	0.642	1.0
Self vs. Tutor 2	1.0	1.0
Peer vs. Tutor 2	1.0	1.0
Tutor 1 vs. Tutor 2	1.0	1.0

Of the sixty-six plastic teeth assessed, the two tutors agreed on their grading for sixty of the prepared teeth.

Self-assessment scoring aligned with the tutor-assessment scoring for forty-three out of sixty-two of the prepared teeth, only slightly lower than the agreement between peer-assessment and tutor-assessment (thirty-nine of the fifty-four prepared teeth).

When not in agreement with the tutor scores, peers gave scores lower than that of the tutor in ten of the fifty-four prepared teeth and higher than the tutor in five of the fifty-four teeth prepared. Similarly, self-assessment was lower than that of the tutor in thirteen of the sixty-two teeth scored, and higher in six of the sixty-two scores given.

When comparing self-assessment with peer-assessment, the scores were in agreement for the assessment of twenty-nine out of forty-six of the prepared teeth. Peer-assessment scores were lower than self-assessment scores in seven out of forty-six of the evaluated teeth and self-assessment scores were lower than peer-assessment scores in ten out of forty-six of the evaluations.

In summary, it could be said that course tutors were in agreement and that peer- and self-assessment appeared to agree with the tutor-assessment two-thirds of the time. When they were not in agreement with the tutor, students tended to underestimate themselves and their peers.

Discussion

Although no statistical significance was seen in this study, this could be due to the small sample size.

In the Edinburgh Dental Institute, the pre-clinical skills courses form part of both the undergraduate and postgraduate programmes. The undergraduate programme was chosen for this project due to its structured format and the number of students being higher than that of individual postgraduate programmes. Although the number of participating students was quite low, the number of prepared plastic teeth which were included in the study per student were increased to try to compensate for this deficit.

In total sixty-six prepared plastic teeth were assessed. The prepared teeth were labelled with a code which included the university student code. Whilst an effort was made to randomise which student assessed which peer, the small group made it likely that the students knew who they were assessing. This, along with the small sample size, were some of the limitations of this study. Further limitations include the inexperience of the students with using the descriptors; with practice and training the results may have been different.

One of the patterns in the data collected which was consistent with the findings of previous studies was the agreement between tutor-assessment and peer-assessment.[15,19]

It has been said that under-performing students are more likely to over-estimate their performance and competent students are able to accurately assess their performance.[7] Therefore, these students who assessed their performance accurately (70% of the time) and when not in agreement with their tutor, under-estimated themselves are likely to be competent in their field.[9,10,20–24]

The level of agreement between peer- and self-assessment was over 60% in this study and this added to their value as a positive indicator of their competency to assess, whether it be themselves or their peers.[11,25]

One study mentioned that if there is a lag time between the task taking place and the assessment of the task it can lead to less agreement between assessors. If this study had taken place on the same day as the session in which the plastic teeth were prepared, the agreement in results might have been higher.[6]

As students become more confident in their competence to accurately assess themselves and their peers, it will mean more frequent feedback could be provided and this will enable more dynamic learning as a whole for the students and with less effort on the part of the tutor.[11,21]

The value of self-assessment, once recognised, makes the process less of a hassle in the eyes of the students, so reporting the findings back to students could have a positive impact on their attitude towards self- and peer assessment.[26]

There is scope for repetition of the study with the inclusion of similar courses in other universities to increase the number of participating students.

Conclusion

Self- and peer-assessment is an important method of assessment which can give similar results to that of tutor-assessment. These skills of assessment will enable a dental student to progress through their career as a self-reflective dental professional who is able to give valuable feedback to their colleagues.

Students tend to underestimate themselves and their colleagues, while tutors assess with good inter-examiner reliability.

With experience and frequent exposure to these methods of assessment, the results and feedback obtained could add value to the learner's experience.

Abbreviations

BSc OHS	Bachelor of Science in Oral Health Science
CMVM EREC	College of Medicine and Veterinary Medicine Educational Research Ethics Committee
C	competent
NC	not competent
UG	Undergraduate Graduate
WHO CPITN	World Health Organisation Community Periodontal Index of Treatment Needs

Declarations

Ethics approval and consent for publication

Ethical approval was obtained from The University of Edinburgh, CMVM EREC (College of Medicine and Veterinary Medicine Educational Research Ethics Committee).

Students were recruited by distributing information regarding the study a week prior to the data collection sessions, and written consent forms were completed on the day of the data collection. Students consented to participate and for publication.

Authors' contributions

LAM designed the study, collected the data, analysed the data and wrote up the paper. ABN and JF supervised the study, assisted in statistical analyses and write up.

References

1. Westberg J, Jason H. Fostering learners' reflection and self-assessment. Fam Med. 1994;26(5):278–82.

2. Sullivan ME, Hitchcock MA, Dunnington GL. Peer and self assessment during problem-based tutorials. Am J Surg. 1999;177(3):266–9.

3. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR, Collaborators IC. The role of assessment in competency-based medical education. Med teach. 2010;32(8):676–82.

4. Norcini J, Holmboe E, Hawkins R. Evaluation challenges in the era of outcomes-based education. In: Practical guide to the evaluation of clinical competence. Philadelphia: Mosby/Elsevier; 2008. pp. 1–9.

5. Dochy F, Segers M, Sluijsmans D. The use of self-, peer and co-assessment in higher education: A review. Stud High Educ. 1999;24(3):331–50.

6. Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. Adv Health Sci Educ. 2002;7(1):63–80.

7. Falchikov N, Boud D. Student self-assessment in higher education: A meta-analysis. Rev Educ Res. 1989;59(4):395–430.

8. Barnsley L, Lyon PM, Ralston SJ, Hibbert EJ, Cunningham I, Gordon FC, et al. Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. Med Educ. 2004;38(4):358–67.

9. Hodges B, Regehr G, Martin D. Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. Acad Med. 2001;76(10):87-S9.

10. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol. 1999;77(6):1121.

11. Topping K. Peer assessment between students in colleges and universities. Rev Educ Res. 1998;68(3):249–76.

12. Falchikov N, Goldfinch J. Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. Rev Educ Res. 2000;70(3):287–322.

13. Lurie SJ, Nofziger AC, Meldrum S, Mooney C, Epstein RM. Temporal and group-related trends in peer assessment amongst medical students. Med Educ. 2006;40(9):840–7.

14. Satterthwaite J, Grey N. Peer-group assessment of pre-clinical operative skills in restorative dentistry and comparison with experienced assessors. Eur J Dent Educ. 2008;12(2):99–102.

15. Evans AW, Leeson R, Petrie A. Reliability of peer and self-assessment scores compared with trainers' scores following third molar surgery. Med Educ. 2007;41(9):866–72.

16. English R, Brookes ST, Avery K, Blazeby JM, Ben-Shlomo Y. The effectiveness and reliability of peer-marking in first-year medical students. Med Educ. 2006;40(10):965–72.

17. Taylor C, Grey N, Satterthwaite J. A comparison of grades awarded by peer assessment, faculty and a digital scanning device in a pre-clinical operative skills course. Eur J Dent Educ 2013;17(1).

18. Manogue M, Brown G, Foster H. Clinical assessment of dental students: values and practices of teachers in restorative dentistry. Med Educ. 2001;35(4):364–70.

19. Foley JI, Richardson GL, Drummie J. Agreement Among Dental Students, Peer Assessors, and Tutor in Assessing Students' Competence in Preclinical Skills. J Dent Educ. 2015;79(11):1320–4.

20. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA. 2006;296(9):1094–102.

21. Boud D. The role of self-assessment in student grading. Assess Eval High Educ. 1989;14(1):20–30.

22. Hacker DJ, Bol L, Horgan DD, Rakow EA. Test prediction and performance in a classroom context. J Educ Psychol. 2000;92(1):160.

23. Fitzgerald JT, White CB, Gruppen LD. A longitudinal study of self-assessment accuracy. Med Educ. 2003;37(7):645–9.

24. Langendyk V. Not knowing that they do not know: self-assessment accuracy of third-year medical students. Med Educ. 2006;40(2):173–9.

25. Boud D. Enhancing learning through self-assessment. 2013. At: https://www.researchgate.net/publication/239969990_Enhancing_Learning_Through_Self-Assessment. Accessed: November 27, 2017.

26. Sargeant J, Eva KW, Armson H, Chesluk B, Dornan T, Holmboe E, et al. Features of assessment learners use to make informed self-assessments of clinical performance. Med Educ. 2011;45(6):636–47.