

Caries Removal by First-Year Dental Students: A Multisource Competency Assessment Strategy for Reflective Practice

Tracy L. de Peralta, DMD, PhD, MClEd; Vidya Ramaswamy, PhD;
Elisabeta Karl, DDS, PhD, MS; Elizabeth Van Tubergen, DDS, PhD;
Mary Ellen McLean, DDS; Mark Fitzgerald, DDS, MS

Abstract: Multisource assessment (MSA) uses multiple assessors to provide feedback. Little is known about the validity of using MSA feedback for improving students' ability to self-assess in a preclinical environment. Therefore, the aim of this study was to measure the validity of using a defined reflective process involving an MSA tool for building skill in dental students' self-evaluation of caries excavation on extracted teeth. As part of this process, 104 first-year students at one U.S. dental school used a self-generated study plan (SGSP) for structured reflection on MSA feedback during the 2013-14 academic year. Interrater agreement, determined through calculation of percentage-agreements in scoring, was measured among three assessor groups (self-, peer, and expert assessors) in formative assessment and between two assessor groups (self- and expert assessors) in summative assessment two weeks apart, allowing for reflective practice and completion of an SGSP between assessments. Validity for improving self-assessment was determined by measuring significance in positive shifts of agreement between self- and expert assessors. The results showed that interrater agreement between the self- and expert assessors increased significantly: from a 28% agreement in formative assessment to a 60% agreement in summative assessment. Significance in percentage shifts between assessments was demonstrated with a McNemar score of 0.26 ($p < 0.001$). These results suggest that the described MSA tool and reflective process in an SGSP may be valid methods for improving skill in student self-evaluation of competence in caries excavation on extracted teeth.

Dr. de Peralta is Clinical Associate Professor, Department of Cariology, Restorative Sciences, Endodontics, and Director of Curriculum and Assessment Integration, School of Dentistry, University of Michigan; Dr. Ramaswamy is Associate Director of Curriculum and Program Evaluation, Office of Academic Affairs, School of Dentistry, University of Michigan; Dr. Karl is Clinical Assistant Professor, Department of Cariology, Restorative Sciences, and Endodontics, School of Dentistry, University of Michigan; Dr. Van Tubergen is Clinical Assistant Professor, Department of Cariology, Restorative Sciences, and Endodontics, School of Dentistry, University of Michigan; Dr. McLean is Clinical Associate Professor, Department of Cariology, Restorative Sciences, and Endodontics, School of Dentistry, University of Michigan; and Dr. Fitzgerald is Associate Chair and Associate Professor, Department of Cariology, Restorative Sciences, and Endodontics, School of Dentistry, University of Michigan. Direct correspondence to Dr. Tracy L. de Peralta, Department of Cariology, Restorative Sciences, and Endodontics, School of Dentistry, University of Michigan, 1011 North University Ave., Ann Arbor, MI 48109-1597; 734-764-8979; tdeperal@umich.edu.

Keywords: dental education, operative dentistry, assessment, self-assessment, dental caries, preclinical education

Submitted for publication 9/1/15; accepted 7/12/16

Reflective learning allows students to critically think through experiences and feedback gained in their learning environment. Student-led approaches have been found to promote reflective learning.¹⁻⁵ Such approaches include peer assessment, peer group discussion, and self-assessment.⁶⁻⁸ Optimizing assessment strategies to determine dental student competence in clinical tasks by using student-led approaches has been of interest in dental education over time.⁹

Multisource assessment (MSA) is one way in which student-led approaches have been applied.

This approach involves one or more assessors providing an evaluation of student performance, based on established grading criteria.¹⁰ MSA is useful in formative assessment, a method designed to provide feedback for the purpose of helping students identify problems and improve their skill in clinical procedures based on reflection.¹¹ Summative assessment, by contrast, is generally a comprehensive evaluation of student performance at the end of a unit of study for the purpose of assigning a grade.¹¹ As students improve their skills through reflective practice based on established grading criteria, they

should also improve their understanding of those criteria. Ultimately, reflective practice is necessary in the novice to expert continuum, allowing students not only to improve in demonstrating competence in any set procedure, but also in their self-evaluation of that procedure.¹²

MSA tools are instruments designed to promote reflective learning on feedback provided in formative assessment; however, limited evidence of their adoption for competency assessment in dental education was noted by Shumway and Harden in 2003.¹³ One reason for this may be associated with dental educators' lack of confidence in using student-led assessment tools.¹⁴ Indeed, a lack of trust in this form of assessment may block the adoption of innovative approaches for advancement in clinical education assessment strategies. Therefore, the aim of this study was to measure the validity of using a defined reflective process involving an MSA tool for building skill in dental student self-evaluation of caries excavation on extracted teeth. This process of formative assessment utilizing MSA was facilitated using a novel IT application developed to concurrently manage and integrate independent assessment from multiple assessors. The outcome of this process was a student self-generated study plan (SGSP). Our operational null hypothesis was that the application of MSA and the reflective process of an SGSP would have no effect on development of dental students' skill in self-assessment between formative and summative assessments of the same procedure two weeks apart.

Materials and Methods

This study received exemption status from the University of Michigan Institutional Review Board (#HUM00087187). This analytic observational study involved quantifying the relationship between a structured reflective process on multisource feedback and its effect on ability to self-assess. To accomplish this, student reflection on multisource feedback from three independent assessors—expert assessor (EA), self-assessor (SA), and peer assessor (PA)—in a formative assessment of student performance in a common clinical procedure took place two weeks prior to a summative assessment of the same clinical procedure. For this study, caries removal was used because it is one of the first clinical procedures on which students are evaluated in the curriculum. Measurement of improving ability to self-assess was achieved by comparing movement of SA scoring to-

ward expert “gold standard” assessment. The experts were dentists who had received approximately five additional hours of assessment training specifically in caries removal. These experts served as instructors throughout the course.

Data collection took place in a simulated dental facility with 104 first-year students in their first Clinical Foundations Course of the predoctoral (DDS) program of the 2013-14 academic year. The caries removal component of this course spanned six weeks during three-hour weekly sessions. During this time, manual dexterity (handpiece use and control with and without direct vision), rubber dam isolation, and mixing and handling of bases and liners were addressed in addition to caries removal. This first-semester, 14-week course was taught by 14 faculty members, each acting as an EA, in an 8:1 student to faculty ratio. After the first four-week process of initial skill development, students completed a formative assessment consisting of an MSA involving SA, PA, and EA, ultimately leading to an SGSP. Students were expected to use this SGSP to improve their performance in caries removal over a two-week period prior to summative assessment using the same grading criteria.

Instruction in Caries Removal

The caries removal exercise used in this study took place on extracted teeth processed according to the infection control protocols of the Centers for Disease Control and Prevention (CDC) regarding academic use of extracted teeth in dental laboratories.¹⁵ In preparation for their clinical skill development in caries removal, students viewed a one-hour lecture on caries removal steps and a series of six videos for guidance on clinical procedures in caries removal. This guidance included how to mount carious extracted teeth in plaster cups, allowing access to natural coronal tooth structure for the excavation procedure and for carious lesions to be kept moist, as well as traditional armamentarium and clinical sequence for this preclinical procedure including turbine handpiece and various carbide and diamond burs for preparation of the convenience form (where appropriate), slow-speed handpiece and carbide round burs (sizes 2, 4, and 6), and spoon excavator, size S1. Since cavity design is influenced by the location and extent of caries present and choice of restorative material, this exercise focused solely on clinical localization of caries, a component of convenience form obtaining access for caries removal, and

appropriate sequence of removing caries. Concepts of cavity design are addressed later in this Clinical Foundations course after the caries removal segment.

Cariou lesions in any part of the coronal region were considered for the excavation exercises. Students were instructed to take the following steps for caries removal, based on guidance found in standard textbooks and consensus recommendations: 1) identification of large cavitated carious lesions (use of radiographic detection was omitted for this exercise on extracted teeth); 2) establishment of convenience form to access caries for removal (the external outline in this case); and 3) removal of soft infected dentin based on operator's tactile sense of soft leathery dentin.¹⁶⁻¹⁸ The removal of infected dentin took place in two steps: 1) intermittent peripheral removal of infected dentin, and 2) central removal of infected dentin directly over the pulp chamber (Figure 1). Should there be risk of pulp exposure with removal of all soft infected dentin centrally, only partial caries removal took place directly over the pulp chamber. Where an indirect pulp capping procedure was indicated in this clinical scenario, the indirect pulp capping procedure was omitted from this competency assessment on caries removal, as assessment of indirect pulp capping procedures take place later in the course. Multiple experiences of caries removal on extracted teeth in the first four-week period of caries removal exercises allowed students to gain

practical skill and knowledge regarding detection of soft leathery caries directly over pulp chambers and identifying when pulp exposure is a risk.

This process allowed time for students to gain multiple experiences based on pre-established grading criteria in this procedure and to receive verbal feedback from experts. Within this learning period, each week included a three-hour faculty-facilitated lab period, as well as non-facilitated after-hours access to the simulation lab to perform multiple caries removal exercises on extracted teeth using the same grading rubric. After this four-week period of initial skill development, the formative assessment in caries removal with MSA took place, with PA and EA providing independent feedback. Because of this more complicated interaction, an IT system for managing formal feedback was used in assessment periods, including both formative and summative types. At this stage, students were asked to reflect on feedback generated in the MSA process by comparing the provided scores and comments with their own SA of the same performance in caries removal.

Students were asked to submit an SGSP 72 hours after the formative assessment period, leaving a two-week period in which three-hour weekly lab sessions continued for reflection before the summative assessment. This summative assessment involved SA and EA only. It did not include PA due to sensitivities students may have with peer involvement in sum-

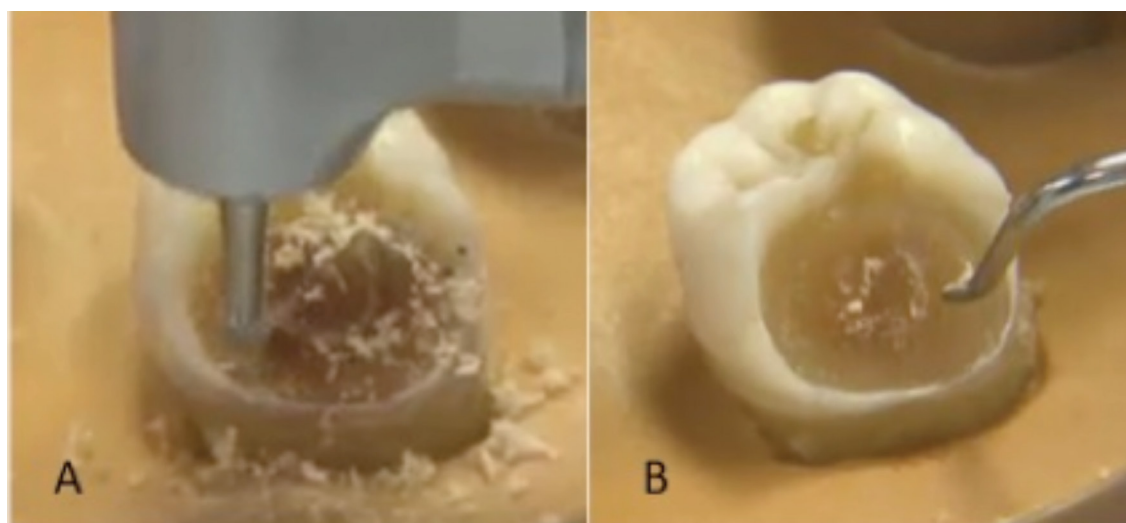


Figure 1. Example of steps for caries removal

Note: Images from 22-minute video on caries removal shown to both students and faculty as a calibration exercise for caries removal. Panel A shows caries removal at the periphery with a slow-speed handpiece no. 2, 4, or 6 round bur. Panel B shows caries removal of soft infected dentin centrally with S1 spoon excavator, with careful attention not to open pulp chamber.

mative scoring. For this reason, data analysis of any shifts in peer scores was not carried out in this study. Furthermore, analysis of any shifts in peer scores from formative assessment to summative assessment was not part of this study; therefore, summative peer scores were not needed.

Assessment Rubric and SGSP

The rubric used was designed for continued use across the curriculum, including both preclinical and clinical learning environments to prepare students for assessments on real patients. Therefore, caries removal on extracted teeth was chosen over using artificial standardization of carious lesions on plastic teeth. This detailed rubric was implemented using an RSTV grading system (R=clinically ideal, S=clinically satisfactory, T=clinically acceptable after modification, V=not clinically acceptable and must be repeated) addressing three specific grading criteria (convenience form, caries removal at the DEJ, and caries removal in dentin) (Table 1). Scores were provided for each grading category, and a summation of scores in each grading category produced a percentage that defined the final score (>60%=passing grade). On the rubric, the grading criteria are clearly

presented for calibrating exercise and for the purpose of standardizing scores among all assessor groups prior to and during both formative and summative assessments. These grading criteria are regularly updated based on evidence each academic year and applied across the curriculum.

Training in interpretation of the grading criteria for the specific preclinical procedure was provided to all assessor groups (SA, PA, and EA) in initial learning sessions. The preclinical lab was arranged in rows of 16 students who were assigned to a team of two faculty members in order to encourage appropriate interpretation of the grading criteria for caries removal. Before the practical component began, a series of six videos were viewed by faculty members and students as a calibration exercise based on grading criteria prior to beginning caries removal on extracted teeth. These videos were available to both groups throughout this period concerning caries removal to reinforce both faculty members' and students' understanding of the grading criteria for this clinical skill. As each student also acted as a PA in the formative assessment period, these four weeks allowed for extensive calibration in which the rubric provided standardization and guidance for calibration on the rubric. Figure 2 shows an ideal student

Table 1. Grading rubric for clinical competence in caries removal

Grade	External Outline Form (Convenience Form)	Caries Removal at DEJ	Caries Removal in Dentin
R Clinically Ideal	External outline appropriately sized for extent of caries (30)	No caries at DEJ (35)	No overcutting Small amount of caries left in deepest, most axial or pulpal part of preparation OK if identified (35)
S Clinically Acceptable	External outline slightly over/underextended for extent of caries (25)	Isolated areas of caries remaining at DEJ (<10%) in non-critical area (30)	Slightly too much caries left in deepest, most axial or pulpal part of preparation Removal of slightly too much tooth structure in deepest area (30)
T Clinically Acceptable, Modifications Preferred	External outline moderately over/underextended for extent of caries (10)	Moderate areas of stain remaining at DEJ (10-50% of DEJ) in non-critical area (15)	Moderate amount of caries left in deepest, most axial or pulpal part of preparation Excessive removal of tooth structure or caries in deepest area (15)
V Clinically Unacceptable: Repeat Procedure or Major Modifications Necessary	External outline severely over/underextended for extent of caries (0)	Extensive areas of stain remaining at DEJ (>50%) Stain at DEJ in critical area Caries remaining at DEJ (0)	Extensive amount of caries left in deepest, most axial or pulpal part of preparation All caries removed too soon or severe overcutting of tooth structure in deepest area (0)

Note: Points available for each assessment appear in parentheses below criteria.

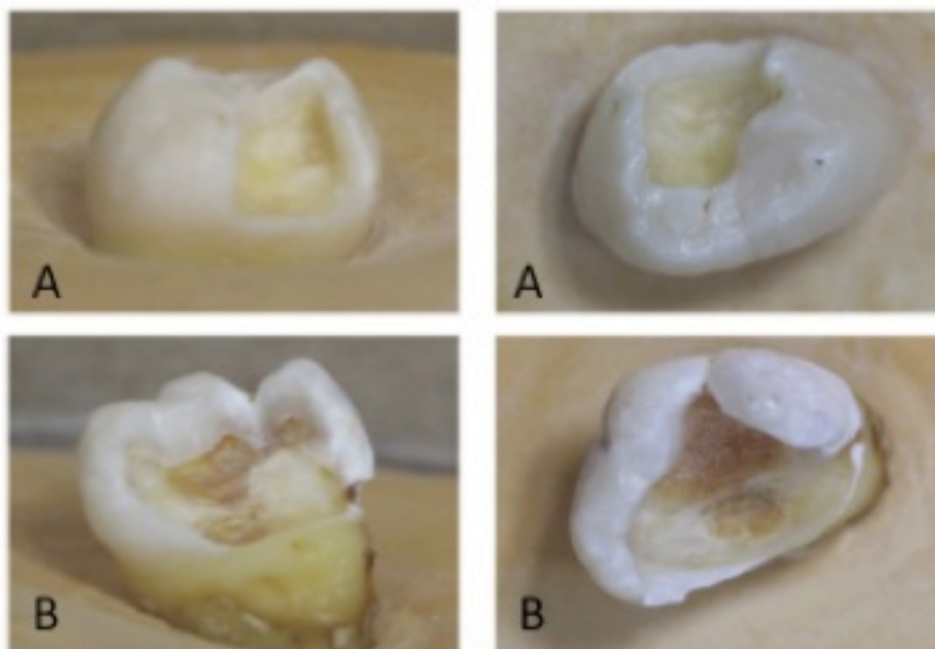


Figure 2. Examples of clinical outcomes based on grading criteria

Note: Panels A, at two angles, show a student's ideal clinical outcome in which an R was granted on all three grading criteria according to the rubric in Table 1. Panels B, at two angles, show a student's failed clinical outcome, in which a V was granted for caries removal at the DEJ and in dentin. Examples of S and T are not shown because they generally require tactile assessment.

outcome and a student failure based on the grading criteria applied from the rubric in Table 1.

The SGSP offered an opportunity for self-assessment of skill development overall, as students aimed to move their independent self-assessment scores closer to that of the experts. A suggested format, including criteria in which improvement was needed and an action plan to improve skills, was provided to the cohort (Table 2). Students' SGSPs were submitted online 72 hours after the formative assessment. Students were also advised to bring their SGSPs to the following lab sessions to share with peers and experts for further feedback and guidance.

Data Collection and Analysis

All assessment scores and feedback were collected with a Google form application allowing real-time score and feedback collection during the assessment period. The Google form is an application built and run on the Google Script platform with open source code developed internally at the University of Michigan School of Dentistry. This code was developed specifically for this assessment; a similar code has been developed internally for use in other

assessments. This IT application is a web-based form requiring each assessor to access a computer with a Google Chrome Browser. Minimal training was needed to access and run this IT application because the application appeared as a simple online survey following student or faculty log-in with their unique username and password. The training to run the Google form took place prior to the start of the assessment process, taking approximately ten minutes. This IT application was developed to automatically identify assessor types as scores were entered independently. Using this system, immediately after the assessment period, a compilation of scores and feedback for each student was individually sent to each in an email summary. This process provided immediate opportunity for students to reflect on their scores and feedback.

Interrater agreement was determined through calculation of percentage agreements in scoring between SA and EA for both formative and summative assessments. Interrater agreement is considered to be good when a high proportion of ratees receive the same score, as was anticipated to be a possibility in this study.¹⁹ Descriptive data were collected to

Table 2. Template and example of student self-generated study plan (SGSP)

Grading Criteria Focus	Timeline	Exercise Description
Appropriate convenience form	10/22 5 pm (after Dent 519 lab session) 10/24 3:00 pm	I will mount more natural teeth into stone. Once these teeth set into the stone, I will practice creating appropriately sized external outlines. I will do this because I realize that shape and depth of a caries lesion will vary dramatically from tooth to tooth. By practicing on multiple teeth, I am certain I will become more aware of when I have actually overextended my external outline. Also, to prevent myself from overextending my external outline, I will remove the tooth structures very slowly and will more constantly check to see if the external outline form is adequate for the corresponding caries lesion.
Caries removal at DEJ	10/22 5:30 pm (after I've practiced creating many correct convenience forms) 10/24 4:00 pm	I will first use the high-speed handpiece to make any last adjustments to the convenience form. Then, I will practice removing caries at the DEJ. Furthermore, I will use the spoon excavator to feel the different tooth surfaces (like infected dentin, sound dentin, enamel, etc.) such that I can calibrate how each feels.
Caries removal in dentin	10/22 7:00 pm 10/24 4:00 pm	During the caries removal formative practical, I was uncertain if I had exposed the pulp. Thus, during these practice sessions I've scheduled for myself, I will deliberately drill into the pulp chamber and the pulp canal to visually see how the connective tissue differs from the rest of the tooth with respect to color and texture. I will use the slow-speed handpiece to calibrate how soft this tissue is compared to other tooth structures. During this step, I will also try to use the spoon excavator to help determine whether all the infected dentin has been removed.

determine any change in grade classification between formative and summative assessments for all assessor groups. Interrater agreement was measured among three assessor groups (SA, PA, and EA) in the formative assessment and between two assessor groups (SA and EA) in the summative assessment for the same cohort of students two weeks apart. Validity of using the described reflective process for improvement of skill in self-assessment was determined by measuring significance of any positive shifts in assessor agreement with the McNemar's test applied to complete assessments, where significance was determined for $p \leq 0.05$.

Results

Whereas 104 students participated in the formative and summative assessments, there were 101 data sets for SA, PA, and EA scores in the formative assessment and 104 data sets for SA and EA scores in the summative assessment. Descriptive data, in the form of percentages of each grade provided by each assessor group (SA, PA, EA), determined positive shifts for self- and expert assessor groups in grading classifications (R=clinically ideal, S=clinically satisfactory, T=clinically acceptable after modification, V=not clinically acceptable and must be repeated)

for each of the grading criteria between assessment periods (Table 3).

Interrater agreement between the SA and EA increased significantly in the scoring from the formative to the summative assessment: from a 28% agreement in the formative assessment to a 60% agreement in the summative assessment, showing a 114% change in agreement between assessment periods. Significance in this percentage shift between assessments was seen with a McNemar score of 0.26 ($p < 0.001$) (Table 4).

Discussion

The aim of this study was to measure the validity of using a defined reflective process involving an MSA tool for building skill in dental student self-evaluation of the same procedure. There were significant findings in this study concerning the relationship between dental students' reflection on MSA feedback in an SGSP and skill in self-evaluation over a two-week period between formative and summative assessments. The results of this study led us to reject our null hypothesis.

We observed a lack of agreement between student assessors (both SA and PA) and EA in the formative assessment, while observing a significant

Table 3. Validity of using multi-assessment tool for improving clinical performance and skill in self-assessment

Grade	Convenience Form			Caries Removal at DEJ			Caries Removal in Dentin		
	Self	Peer	Expert	Self	Peer	Expert	Self	Peer	Expert
Formative									
R	59.6%	84.6%	89.4%	75.0%	82.7%	66.3%	67.3%	76.0%	76.9%
S	35.6%	10.6%	9.6%	22.1%	13.5%	22.1%	27.9%	21.2%	19.2%
T	2.9%	1.0%	1.0%	1.0%	1.0%	6.7%	2.9%	1.9%	1.9%
V	0	1.0%	0	0	1.0%	4.8%	1.0%	0	1.9%
Number	102	101	104	102	102	104	103	103	104
Summative									
R	90.4%	–	95.2%	83.7%	–	77.9%	82.7%	–	82.7%
S	9.6%	–	4.8%	16.3%	–	21.2%	17.3%	–	17.3%
T	0	–	0	0	–	1.0%	0	–	0
V	0	–	0	0	–	0	0	–	0
Number	104		104	104		104	104		104

Note: Positive shifts were noted in percentage of grade type (RSTV) provided for each grading criterion by each assessor group (self, peer, expert) between assessment periods two weeks apart. (See Table 1 for definition of and criteria for each grade.) Total respondents listed for each category. Percentages may not total 100% due to incomplete responses. Incomplete responses were found in less than 3% of student assessor population for the formative assessment only.

Table 4. Interrater agreement

Grading Criteria	Formative	Summative	% Change	McNemar's Test (p-value)
Convenience form	59%	86%	46%	0.27 (p<0.001)
Caries removal at DEJ	59%	71%	20%	0.13 (p=0.05)
Caries removal in dentin	62%	81%	31%	0.19 (p<0.01)
Global score	28%	60%	114%	0.26 (p<0.001)

increase in interrater agreement between SA and EA after the two-week period between assessments. These results indicate that students improved their ability to self-assess when having the opportunity to reflect on independent feedback from multiple sources and apply critical thinking in the process of producing their SGSPs. Interestingly, while there was great discrepancy in assessor agreement for the grading criteria concerning convenience form in the formative assessment, a significant improvement in agreement was seen for convenience form (McNemar score of 0.27; p<0.001). In general, expert scores were higher for these grading criteria, where SA were much harder on scoring compared to EA and PA. Also, while the peer scores tended to be closer to the experts in the criteria looking at convenience form, the peer scores were very high compared to both SA and EA for the removal of caries at the DEJ and were closer to the expert for caries removal in dentin. Furthermore, the experts were overall less satisfied with the students' ability to remove caries at the DEJ than the students were, where peers in the

formative assessment gave higher grades than both SA and EA. Grading criteria of caries removal at the DEJ showed an interrater agreement between SA and EA (McNemar score of 0.13; p=0.05) between assessment periods.

There appeared to be no particular trend in interrater agreement involving peer assessment. Where interrater agreement among SA, PA, and EA was poor for the formative assessment, the benefit of the students' applying critical thinking in their role as PA theoretically added strength to the reflective learning process and may have contributed to the positive shift in interrater agreement for the SA and EA between assessment periods. As appropriate and effective clinical assessment requires assessors to identify and apply clear grading criteria, Redwood et al. found that students' ability to apply assessment criteria and retain fundamental clinical concepts over time improved with ongoing self-assessment and tutor feedback throughout the first year of a dental curriculum.²⁰ This finding supports the theory that reflection on learning experiences and skill-building

in critical thinking assist progression in the novice to expert continuum.¹¹ However, there has been some controversy in the past that students, whether they be SA or PA, can accurately make judgments on clinical outcomes.^{21,22} In acting as a PA in the formative assessment period, students were given the opportunity to apply grading criteria in a constructive and fair manner, offering a different perspective on application of the grading criteria than the perspective they may have in grading their own work. Therefore, the peer assessment element was an important part of this multisource process for reflective learning. This element of the multisource assessment system is distinctive and provided a path for reinforcement of criteria that the SA was initially aiming to achieve.

Student-led learning was found by Hoogenes et al. to be an important curriculum instrument that offered students an environment in which they were free to explore and develop their own cognitive processing for skill acquisition rather than being forced to work in and adapt to framework templates.²³ In student-led learning processes, such as the SGSP implemented in this study, the primary focus is to provide an approach that allows students to take control over their learning process and give them a sense of responsibility for learning.

One limitation of this study was the lack of analysis of qualitative justification for scoring from each assessor group in both the formative and summative assessments. While considering the comments provided for score justification between assessor groups would have given us a richer data set, compliance to provide the requested written justification was low. When asked why this was the case, study participants stated that this score justification was an optional step in the feedback process, and while the opportunity to provide score justification was available, time constraints were often prohibitive in providing this feedback.

Another limitation was the inability to compare peer scores with that of EA and SA in the summative assessment. While this measurement was not necessary according to the aims of the study, it would have been interesting to see if improvement in peer assessment, compared to the gold standard of the EA, took place after a reflective period of MSA using the SGSP. In this study, PA scoring functioned purely as a reflective exercise for student assessors. Two prior studies found that peer assessment was a valued student-led approach to assessment strategies,^{23,24} but more evidence is needed to support this important assessment tool for future assessment strategies.

A third limitation of this study was the design of the rubric, which some found confusing at first glance and needed practice in applying. One particular field that often needed explanation was the S category (clinically acceptable) of the grading criteria for external outline form. The notation “slightly over/underextended for extent of caries” was difficult to interpret for some and would have been better defined as “slightly over- or underextended for extent of caries.”

Conclusion

This study found that students’ self-assessment moved closer to the expert’s gold standard after they were given the opportunity to reflect on independent feedback from three assessor types (SA, PA, EA) in an SGSP. Interrater agreement showed significant positive shifts from formative assessment to summative assessment for these first-year dental students. These findings support the validity of using formative MSA feedback and an SGSP for the purpose of improving first-year dental students’ self-assessment. The results of this study based on application of a novel IT system for assessment data collection and summary promote advancement in clinical assessment development and students’ reflective learning process associated with multisource assessment.

Acknowledgments

We would like to thank Ms. Kari Gregerson for her technical support in this study. We would also like to thank Mr. Sean Meyer for his work in developing the open source code on a Google Script platform for collection of data. Furthermore, we would like to acknowledge that this research has been made possible by a generous contribution from the Roy H. Roberts Gift Fund at the University of Michigan.

REFERENCES

1. Knobe M, Munker R, Sellei RM, et al. Peer teaching: a randomized controlled trial using student-teachers to teach musculoskeletal ultrasound. *Med Educ* 2010;44:148-55.
2. Kibble J, Hansen PA, Nelson L. Use of modified SOAP notes and peer-led small-group discussion in a medical physiology course: addressing the hidden curriculum. *Adv Physiol Educ* 2006;30:230-6.
3. Hulsman RL, Harmsen AB, Fabriek M. Reflective teaching of medical communication skills with DiViDU: assessing the level of student reflection on recorded consultations with simulated patients. *Patient Educ Couns* 2009;74:142-9.

4. Goldsmith M, Stewart L, Ferguson L. Peer learning partnerships: an innovative strategy to enhance skill acquisition in nursing students. *Nurse Educ Today* 2005;26:123-30.
5. Dannefer EF, Henson LC, Bierer SB, et al. Peer assessment of professional competence. *Med Educ* 2005;39:713-22.
6. Botelho MG, O'Donnell D. Assessment of the use of problem-oriented, small-group discussion for learning of a fixed prosthetic, simulated laboratory course. *Br Dent J* 2001;191:630-6.
7. Chaves JF, Baker CM, Chaves JA, et al. Self, peer, and tutor assessments of MSN competencies using the PBL-evaluator. *J Nurse Educ* 2006;45:25-31.
8. Evans AW, Leeson MA, Petrie A. Reliability of peer and self-assessment scores compared with trainers' scores following molar surgery. *Med Educ* 2007;41:866-72.
9. McHarg J, Kay EJ. Designing a dental curriculum for the twenty-first century. *Br Dent J* 2009;206:493-7.
10. Kramer GA, Albino JEN, Andrieu SC, et al. Dental student assessment toolbox. *J Dent Educ* 2009;73(1):12-34.
11. Anziani H, Durham J, Moore U. The relationship between formative and summative assessment of undergraduates in oral surgery. *Eur J Dent Educ* 2008;12:233-8.
12. Daley BJ. Novice to expert: an exploration of how professionals learn. *Adult Educ Q* 1999;49:133-47.
13. Shumway JM, Harden RM. The assessment of learning outcomes for the competent and reflective physician. AMEE guide no. 25. *Med Teach* 2003;25(6):569-84.
14. Albino JEN, Young SK, Neumann LM, et al. Assessing dental students' competence: best practice recommendations in the performance assessment literature and investigation of current practices in predoctoral dental education. *J Dent Educ* 2008;72(7):823-6.
15. Division of Oral Health, Centers for Disease Control and Prevention. Infection control frequently asked questions: extracted teeth. 2016. At: www.cdc.gov/oralhealth/infectioncontrol/faq/extracted_teeth.htm. Accessed 17 June 2016.
16. Heyman H, Swift E, Ritter A, et al. *Sturdevant's art of science of operative dentistry*. 6th ed. St. Louis, MO: Mosby, 2012.
17. Summit JB, Robins JW, Hilton JT, et al. *Fundamentals of operative dentistry: a contemporary approach*. 3rd ed. Chicago: Quintessence, 2006.
18. Schwendicke F, Frencken JE, Bjorndal L, et al. Managing carious lesions: consensus recommendations on carious tissue removal. *Adv Dent Res* 2016;28(2):58-67.
19. Graham M, Milanowski A, Miller J. Measuring and promoting interrater agreement of teacher and principal performance settings. At: files.eric.ed.gov/fulltext/ED532068.pdf. Accessed 5 Dec. 2014.
20. Redwood C, Winning T, Lekkas D, et al. Improving clinical assessment: evaluating students' ability to identify and apply clinical criteria. *Eur J Dent Educ* 2010;14:136-44.
21. Eva KW, Cunnington JP, Reiter HI, et al. How can I know what I don't know? Poor self-assessment in a well-defined domain. *Adv Health Sci Educ Theory Pract* 2004;9:211-24.
22. Dochy F, Segers M, Sluigsmans D. The use of self-, peer, and co-assessment in higher education: a review. *Stud Higher Educ* 1999;24:331-50.
23. Hoogenes J, Mironova P, Safir O, et al. Student-led learning: a new teaching paradigm for surgical skills. *Am J Surgery* 2015;209:107-14.
24. Bostick G, Hall M. Novel clinical learning from a student-led clinic. *Clin Teach* 2014;11:512-5.