

# Effect of diagnostic score reporting following a structured clinical assessment of dental hygiene student performance

Alix Clarke\*, MSc; Hollis Lai<sup>§</sup>, PhD; Alexandra DE Sheppard\*, RDH, BA, MEd; Minn N Yoon<sup>Δ</sup>, PhD

## PRACTICAL IMPLICATIONS OF THIS RESEARCH

- The ability to accurately self-assess is an essential skill for practising dental hygienists.
- Diagnostic score reporting may help develop this skill during dental hygiene education so it can carry over into professional practice.

## ABSTRACT

**Background:** Diagnostic score reporting is one method of providing feedback to all students following a structured clinical assessment but its effect on learning has not been studied. The objective of this study was to assess the impact of this feedback on student reflection and performance following a dental hygiene assessment. **Methods:** In 2016, dental hygiene students at the University of Alberta participated in a mock structured clinical assessment during which they were randomly assigned to receive a diagnostic score report (intervention group) or an overall percentage grade of performance (control group). The students later reflected upon their performance and took their regularly scheduled structured clinical assessment. Reflections underwent content analysis by diagnostic domains (eliciting essential information, effective communication, client-centred care, and interpreting findings). Results were analysed for group differences. **Results:** Students performed best on eliciting essential information (92%) and poorest on interpreting findings (42%). The intervention group was more likely to view interpreting findings as a weakness,  $p = 0.007$ , while the control group was more likely to view eliciting essential information as a weakness,  $p = 0.04$ . No differences were found on the actual assessment scores,  $p > 0.05$ . **Discussion:** Students who received diagnostic score reporting appeared to reflect more accurately upon their weaknesses. However, this knowledge did not translate into improved performance. Modifications and enhancements to the report may be necessary before an effect on performance will be seen. **Conclusion:** Diagnostic score reporting is a promising feedback method that may aid student reflection. More research is needed to determine if these reports can improve performance.

## RÉSUMÉ

**Contexte :** Le suivi de la notation des diagnostics est une des méthodes utilisées pour fournir de la rétroaction aux étudiants à la suite d'une évaluation clinique structurée, mais ses effets sur l'apprentissage n'ont pas été étudiés. La présente étude visait à évaluer l'effet de cette rétroaction sur la réflexion et la performance des étudiants à la suite d'une évaluation en hygiène dentaire. **Méthodologie :** Les étudiants en hygiène dentaire de l'Université de l'Alberta ont participé à une évaluation clinique structurée fictive pendant laquelle ils étaient désignés de façon aléatoire pour recevoir un suivi de la notation des diagnostics (groupe d'intervention) ou une note globale en pourcentage de leur performance (groupe témoin). Les étudiants ont plus tard réfléchi à leur performance et ont fait leur évaluation clinique structurée déjà à l'horaire. Une analyse de contenu a été effectuée sur les réflexions selon les domaines de diagnostics (obtention de renseignements essentiels, communication efficace, prestation de soins axés sur le client et interprétation des constatations). Les résultats ont été analysés pour déterminer les différences entre les groupes. **Résultats :** Les étudiants ont le mieux réussi sur l'obtention d'information essentielle (92 %) et ont le moins bien réussi sur l'interprétation des constatations (42 %). Le groupe d'intervention était plus susceptible de réfléchir à l'interprétation des constatations en tant que faiblesse,  $p = 0,007$ , alors que le groupe témoin était plus susceptible de réfléchir à l'obtention de l'information essentielle en tant que faiblesse,  $p = 0,04$ . Aucune différence n'a été trouvée sur les notations actuelles de l'évaluation,  $p > 0,05$ . **Discussion :** Les étudiants qui ont reçu un suivi de la notation des diagnostics semblaient réfléchir plus précisément sur leurs faiblesses. Cependant, cette connaissance ne s'est pas traduite par une performance améliorée. Des modifications et des améliorations du suivi peuvent être nécessaires avant qu'un effet sur la performance soit constaté. **Conclusion :** Le suivi de la notation des diagnostics présente une méthode de rétroaction prometteuse qui pourrait aider à la réflexion des étudiants. D'autres recherches sont nécessaires pour déterminer si ces rapports peuvent améliorer la performance.

**Keywords:** dental education; dental hygiene; diagnostic score reporting; feedback (learning); formative feedback; objective structured clinical examination (OSCE)

**CDHA Research Agenda category:** capacity building of the profession

\*Statistical associate, University of Calgary, Cummings School of Medicine, Calgary, AB, Canada

<sup>§</sup>Associate professor and assistant dean of education quality and accreditation, University of Alberta, Faculty of Medicine and Dentistry, Edmonton, AB, Canada

\*Clinical professor, University of Alberta, School of Dentistry, Edmonton, AB, Canada

<sup>Δ</sup>Associate professor, University of Alberta, School of Dentistry, Edmonton, AB, Canada

Correspondence: Alix Clarke; alix.clarke1@ucalgary.ca

Manuscript submitted 16 March 2020; revised 20 August 2020; accepted 31 August 2020

©2021 Canadian Dental Hygienists Association

## INTRODUCTION

This article is 1 of 2 papers published in this issue that report on the development, implementation, and evaluation of a diagnostic score reporting framework for structured clinical assessments in dental hygiene. A literature review describes an evidence-based methodology for developing valid reports that summarize student performance (based on diagnostic domains) and for implementing that framework within a dental hygiene skills-based assessment.<sup>1</sup> This article reports the evaluation component, specifically how students responded to the feedback, as feedback must be both valid and useful to be considered of quality.

Feedback is considered an essential pedagogical tool in higher education.<sup>2,3</sup> While assessments evaluate student performance, feedback provides students with detailed information to improve future performances.<sup>2,4,5</sup> Many educators believe students must receive feedback throughout their education for effective learning to take place.

Kolb<sup>6</sup> describes learning as a cycle. The process begins with experience, which leads to reflection, observations of others, analysis, conceptualization, and finally modified behaviours. Feedback can play an integral role in facilitating this cycle.<sup>3,7,8</sup> Feedback encourages individuals to use their experiences to make improvements, by identifying when performance has deviated from expectations (i.e., starts the cycle).<sup>7,8</sup> Feedback also guides students through the learning process by improving the accuracy of self-assessments and suggesting where reflection and analysis should be focused (i.e., accelerates the cycle).<sup>3,9,10</sup> Reviews have found that externally provided feedback (such as written or verbal information on, or assessment of, a student's performance or technique) generally leads to improved performance.<sup>11-13</sup> However, not all feedback will have a positive impact, with guidelines recommending timely, clear and concise, task-specific information from an authoritative figure, given with the intent to help the student improve.<sup>4,5,11</sup>

Despite its established importance, feedback is frequently reported as inadequate across the health disciplines.<sup>2,14</sup> Structured clinical assessments (SCAs) capture detailed information on students' knowledge, skills, and abilities (i.e., their competence), presenting key opportunities to provide feedback.<sup>14-17</sup> The most common SCA is the multistation objective structured clinical examination (OSCE) seen in medical education<sup>18,19</sup>; single client assessments, which often focus on interpersonal communication, are also common in nursing and dental hygiene.<sup>20,21</sup> These assessments are typically high-stakes, and they attempt to mirror authentic clinical practice in a standardized manner, such as by using trained actors and predetermined grading checklists.<sup>18-21</sup>

Research suggests that the majority of students will review feedback offered following an SCA.<sup>22</sup> However, feedback is often limited in these types of assessments. Traditional methods, such as providing the student with their results by test item, may be precluded by concerns

over test security.<sup>23,24</sup> Other methods, such as having the examiner provide immediate verbal or written comments within the context of the assessment, significantly increase the time demands placed on instructors and often lead to rushed, incomplete feedback.<sup>14,25</sup> A lack of instructor time is frequently cited as a key issue precluding the provision of feedback.<sup>9,26,27</sup>

Diagnostic score reporting (DSR) is a possible means of providing feedback to students following an SCA. DSR provides test-takers with information on global performance, performance by domains, relative standings to professional standards and/or peers, and specific suggestions for improving performance.<sup>28-30</sup> Domains reflect the specific areas of knowledge, skill or ability that the examination intends to capture, defined by professional standards and competencies, so that student strengths and weaknesses can be readily identified.<sup>28,29,31</sup> Domain scores are determined from a subset of relevant test items and provide additional information to the student beyond a single summary score, including information on how to improve within those areas.<sup>28,32,33</sup> DSR does not reveal the actual test items to students, and the content of the reports can be pre-established and delivered efficiently online, overcoming the major barriers to feedback for SCAs. DSR has been largely confined to nationwide testing programs, and much of the literature has focused on reporting features such as usability and interpretability.<sup>28,29,31-34</sup> The effect of DSR on student outcomes has not been well studied.

The aim of this study was to assess the impact of DSR following an SCA on student-level outcomes, specifically reflection and performance.

## METHODS

Ethics approval for this study was obtained from the University of Alberta Research Ethics Office (Pro00062297).

Our companion paper describes the literature-based development and validation process for implementing DSR within a dental hygiene SCA at the University of Alberta in 2016.<sup>1</sup>

### The dental hygiene structured clinical assessment

The dental hygiene SCA was a comprehensive single-client assessment requiring students to establish a rapport, conduct a full health and dental history, and identify any risk factors contraindicating or requiring modifications to dental hygiene therapy.<sup>35</sup> The client was portrayed by a standardized patient (a trained actor), who received prior information on the client's demographics, medical health and background, and oral health status and beliefs. The SCA was graded by clinical instructors using a checklist of observable items, such as questions the student must ask and conclusions they should make during the encounter—marked yes or no—and rating scales to assess global skills such as organization and communication. There were 5 clinical instructors who marked an average of 8 students. No formal calibration was conducted, although instructors had previous SCA experience and met

with the clinical professor before and after the assessment to familiarize themselves with the checklist and address any questions or concerns.

### The diagnostic score report

DSR for the dental hygiene SCA included overall scores on the assessment and a breakdown of scores by 4 skill-based domains: 1) effective communication; 2) client-centred care; 3) eliciting essential information; and 4) interpreting findings. The 4 domains were established using the Canadian dental hygiene entry-to-practice competencies<sup>36</sup> for guidance. Test items were then mapped to an appropriate domain so students could receive a domain score. The reports also contained peer comparisons and information on how to improve performance within each of the 4 skills. Careful consideration was given to the reporting features (e.g., esthetics and language) in keeping with best practice recommendations.<sup>1</sup> Reports were provided to students through an online portal.

An assessment blueprint, which shows how test items and competencies were mapped to domains, as well as screenshots of the actual report, is presented in the literature review.<sup>1</sup>

### Study design

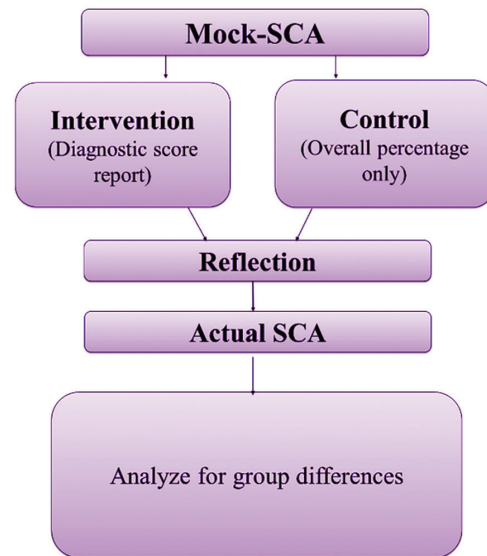
An experimental design was used to evaluate the effect of DSR on student-level outcomes. Thirty-nine dental hygiene students scheduled to take the SCA in 2016 were invited to participate in a mock SCA, a practice assessment prior to their final SCA, where they were randomly assigned to either the intervention or control group. The intervention group received DSR after the mock SCA, while the control group received one overall percentage grade of performance. After receiving their results, all students were asked to reflect upon their mock SCA performance and later completed their regularly scheduled SCA. Thus the effect of DSR on reflection and performance could be ascertained. Figure 1 illustrates this design.

### The mock structured clinical assessment

The mock SCA covered the same diagnostic domains as the actual SCA, but with different content. The client's demographics, health history, and oral health perceptions and concerns were thus unique, but the skills required for the assessment were identical. Two content experts (AC & AS) developed the mock SCA using subject matter expertise and course resources such as the clinic manual and course textbook.<sup>37</sup> An assessment blueprint was established linking the test items to the diagnostic domains. The reliability of these links was further validated by clinical instructors using a modified Delphi approach to reaching expert consensus.<sup>38</sup>

The mock SCA ran identically to the regularly scheduled SCA (as described above), with standardized patients given their client history prior to the assessment, and 5 clinical instructors grading using a predetermined checklist

Figure 1. Experimental design for evaluation of student-level outcomes



(although standardized clients and instructors were not necessarily the same for each SCA).

### Experimental procedure

The mock SCA took place 12 days before the actual SCA. Both students and graders were blinded to group assignment (DSR or control), and verbal feedback from clinical instructors was prohibited. One day after the mock SCA, students received an email link to either their DSR (intervention group) or their overall percentage of performance (control group). Five days after the mock SCA, students were prompted to reflect upon their performance. Specifically, students were asked to describe what they did well during the mock SCA, recognizing their strengths, and what they could improve upon, identifying their weaknesses. Reflections were coded for their quality and content. The regularly scheduled SCA took place one week later and results were collected for analysis. All students received DSR for the actual SCA once the study data were collected.

### Data coding and analysis

All data analysis was conducted using statistical software STATA 14.<sup>39</sup> Descriptive statistics were reported using means or percentages, where appropriate. For inferential statistics, in cases of violations of assumptions, equivalent non-parametric tests were used in place of parametric ones.

### Analysis of performance

The SCA results were analysed for group differences using linear regression controlling for mock SCA results, according to best practices in test-retest designs.<sup>40</sup>

### Analysis of reflection quality

To determine the quality of the students' reflections, a grading rubric was developed using the University of Alberta Health Sciences Education and Research Commons (HSERC) Interprofessional Reflection Guide.<sup>41</sup> Through an iterative process, the rubric was adjusted until 2 trained raters (AC & MY) could reliably evaluate the reflections; raters were blinded to group assignment. The final rubric dichotomized reflective statements as either low- or high-quality (Appendix A). Exact inter-rater agreement was 84% with a Cohen's Kappa of 0.64, indicating substantial agreement,<sup>42</sup> and any remaining discrepancies were reviewed and coded via consensus. Results were analysed using independent t-tests.

### Analysis of reflection content

Each reflection underwent content analysis.<sup>43</sup> The units of analysis were the diagnostic domains (effective communication, client-centred care, eliciting essential information, and interpreting findings). The rubric for coding reflective statements into the appropriate domains was based on the assessment blueprint (Appendix B). One statement could have represented multiple domains and an "other" category captured reflections that did not fit within the domains. Two blinded researchers (AC & MY) independently coded each reflective statement. Exact inter-rater agreement was 70% with another substantial Cohen's Kappa of 0.64. Discrepancies were reviewed and coded via consensus. Reflection content was analysed using Poisson regression.

Table 1. Structured clinical assessment results: Mean % (SD)

	Control group (n = 20)	DSR group (n = 18)	Total (N = 38)
<b>Mock SCA</b>			
Total	73.67 (7.69)	76.30 (7.89)	74.92 (7.80)
Effective communication	75.28 (13.55)	79.01 (10.18)	77.05 (12.06)
Client-centred care	76 (10.01)	76.67 (10.54)	76.32 (10.13)
Eliciting essential information	91.67 (9.45)	92.59 (10.08)	92.11 (9.63)
Interpreting findings	37.86 (10.65)	46.83 (22.35)	42.11 (17.56)
<b>Actual SCA results</b>			
Total	82.96 (6.60)	82.20 (7.69)	82.60 (7.05)
Effective communication	88.16 (10.23)	82.16 (9.40)	85.32 (10.18)
Client-centred care	83.00 (10.81)	83.33 (9.70)	83.16 (10.16)
Eliciting essential information	92.31 (8.65)	93.59 (8.02)	92.91 (8.27)
Interpreting findings	45.00 (26.72)	52.08 (25.36)	48.36 (25.36)

Note: nothing significant at  $p < 0.05$

## RESULTS

Thirty-eight (38) of the thirty-nine (39) students (97%) participated in the mock SCA, and 37 (95%) submitted a reflection.

### The effect of diagnostic score reporting on performance

The results of the mock SCA and actual SCA are found in Table 1. The average score on the mock SCA was 75%. Students scored best on the domain eliciting essential information, averaging 92%. Students performed worst on interpreting findings, averaging 42%. No significant differences were found between the control group and the DSR group on the mock SCA,  $p > 0.05$ .

The average score on the actual SCA was 83%. Similar trends were found on domain scores, with students performing best on eliciting essential information (93%) and worst on interpreting findings (48%). Controlling for mock SCA scores, there was no significant difference on overall SCA scores between the DSR group and control group,  $p > 0.05$ . There was a non-significant trend that the DSR group performed better than the control group on interpreting findings by an average of 10%,  $p = 0.24$ . There was a borderline significant difference on effective communication scores,  $p = 0.05$ , where the DSR group performed poorer than the control group by 7%. Overall, there was limited evidence to support that DSR improved student performance.

### The effect of diagnostic score reporting on reflection

The quality of the students' reflections is described in Table 2. The average score was 0.40, falling below the midpoint of the scoring rubric, which ranged from 0 to 1. No significant differences between the DSR and control group were found for reflection quality overall or by question (i.e., strengths vs weaknesses),  $p > 0.05$ .

The result of the content analysis is presented in Table 3. The students' reflections fit well within the 4 diagnostic domains, with only 7% of the comments allocated to the "other" category. The majority of the comments focused on eliciting information (33%) and effective communication (29%), with fewer reflective statements on client-centred care (18%) and interpreting findings (13%). The 2 prompts on strengths and weaknesses were analysed separately for group differences.

Table 2. Quality of student reflections: Mean (SD)

	Control (n = 20)	DSR (n = 17)	Total (N = 37)
Strengths	0.5 (0.43)	0.37 (0.45)	0.44 (0.44)
Improvements	0.38 (0.46)	0.34 (0.42)	0.36 (0.44)
Total	0.44 (0.34)	0.36 (0.27)	0.40 (0.31)

Note: nothing significant at  $p < 0.05$

Table 3. Content of student reflections: Frequency (%)

		Effective communication	Client-centred care	Eliciting essential information	Interpreting findings	Other
Strengths	Control	11 (22%)	8 (16%)	21 (40%)	6 (12%)	4 (8%)
	DSR	13 (35%)	7 (19%)	14 (38%)	1 (3%)	2 (5%)
	Combined	24 (28%)	15 (17%)	35 (40%)	7 (8%)	6 (7%)
Weaknesses	Control	15 (35%)	8 (19%)	15 (35%) <sup>a</sup>	2 (5%) <sup>a</sup>	3 (7%)
	DSR	8 (24%)	7 (21%)	4 (12%) <sup>a</sup>	13 (38%) <sup>a</sup>	2 (6%)
	Combined	23 (30%)	15 (19%)	19 (25%)	15 (19%)	5 (6%)
Total	Control	26 (28%)	16 (17%)	36 (39%)	8 (9%)	7 (8%)
	DSR	21 (30%)	14 (20%)	18 (25%)	14 (20%)	4 (6%)
	Combined	47 (29%)	30 (18%)	54 (33%)	22 (13%)	11 (7%)

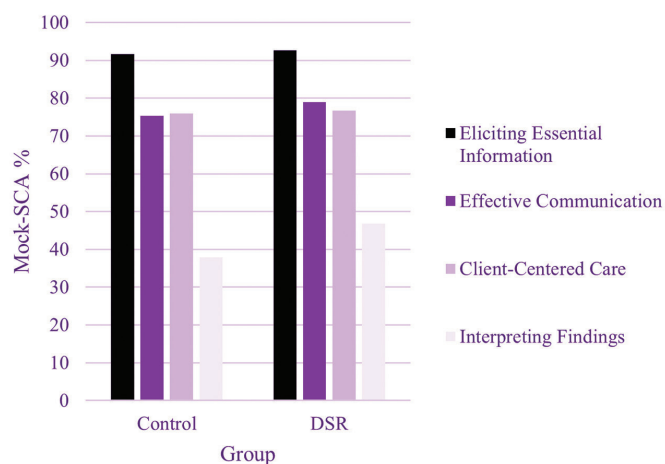
<sup>a</sup>significant at  $p < 0.05$

Percentages reflect row totals and may not sum to 100% due to rounding.

Significant differences were found in the identification of weaknesses. The DSR group had 7.65 times the average number of comments on interpreting findings than the control group,  $p = 0.007$ , and 0.31 times (or 70% fewer) comments on eliciting essential information,  $p = 0.04$ . Figure 2 illustrates that students within both groups performed best on the domain eliciting essential information and worst on interpreting findings. Therefore, the DSR group appeared to reflect more accurately upon their weaknesses.

In regard to identifying strengths, the DSR group was more likely to reflect upon eliciting essential information as a strength compared to the control group (35% vs 22%), and only 1 student in the DSR group reflected upon interpreting findings as a strength compared to 6 students in the control group (3% vs 12%). While these differences were not significant,  $p > 0.05$ , these trends support the claim that DSR resulted in more accurate self-assessments.

Figure 2. Mock structured clinical assessment percentages by domain



## DISCUSSION

DSR is a viable framework for providing feedback to all students following an SCA. DSR describes test results by the underlying domains the test intends to measure and includes resources for making individual-level improvements.<sup>28,29,32,33</sup> Despite the potential of DSR, its effect on student-level outcomes has not been well investigated.

Reflection is a key component of learning.<sup>6,8,44</sup> Through the process of experiencing, reflecting, thinking, and acting,<sup>6</sup> each SCA provides a learning opportunity that could improve future clinical practice. Accurate self-assessment aids this process by helping students to notice when their performance needs improvement, prompting the learning cycle to start,<sup>7,8</sup> and may speed up learning by directing students to focus on key issues, eliminating uncertainty or misdirection.<sup>3</sup> The ability to correctly identify problem areas is one important element of reflection-in-action, a life-long skill essential for health care professionals.<sup>8</sup> Feedback is believed to facilitate this type of reflection<sup>9,45</sup> and yet feedback is often limited following an SCA due to time limitations and concerns over test security. DSR overcomes these issues and this study showed that this feedback mechanism did appear to encourage student reflective capabilities in a positive direction.

Nevertheless, this study did not find that DSR facilitated improved performance. While students were better able to identify their strengths and weaknesses, this awareness did not translate into a behaviour change. These findings suggest a breakdown of the learning cycle between the stages of reflection and action. Ultimately, the goal of feedback is to improve performance,<sup>2,5,8</sup> and reasons for DSR not achieving this goal should be examined.

One possibility is that DSR may need to be provided consistently over a longer time span before it will have any impact. A review of feedback and physician performance showed that studies with longer durations were more likely to find significant effects.<sup>11</sup> Similarly, a review of audit and

feedback found that feedback was more effective when provided more than once.<sup>12</sup> This current study took place over a few weeks and only provided DSR a single time.

Another explanation could be that simply providing resources to students is insufficient to encourage meaningful interactions with that information. Even if students are aware of their weaknesses, they may lack the motivation to improve. Characteristics such as confidence, self-esteem, self-efficacy, personal beliefs, and intrinsic interests can all affect a student's ability to self-regulate and become accountable for their own learning.<sup>46-48</sup> The design of this study made it difficult to determine exactly *how* students interacted with their feedback, and more qualitative research in this area may help pinpoint key motivational factors.

Finally, performance may not have improved because of issues with the feedback itself. Quality feedback should be task specific,<sup>2-5,9,27</sup> and the information provided through DSR may not have been specific enough to alter behaviour. Providing students with the actual test items would be more precise, but for high-stakes SCAs, where test security is a chief concern, this option is not feasible. When actual test items cannot be revealed, providing example test items is a practical alternative.<sup>32</sup> Other options include providing web links to specific sections of textbooks or other online documents,<sup>28,49</sup> or video-based feedback where students view their own performance or exemplars of professional dental hygienists interacting with clients.<sup>3,50,51</sup> The degree of report individualization could also be improved.<sup>28,33,49</sup> Different reporting components could be displayed based on different SCA checklist response patterns or different levels of achievement. Low achievers might benefit from information on how to interpret their scores, while advanced students might benefit from supplemental learning activities.<sup>49</sup> Online DSR can be expanded and modified to better encourage student-level improvements.

A limitation of this study was its small sample size, which limited statistical power and precluded more in-depth analyses. Results may not be generalizable to other health professions. This research focused on a select dental

hygiene student population at a single school within a single term. Furthermore, each dental hygiene student interacted with only a single standardized patient and was graded by one clinical instructor (as compared to an OSCE that uses multiple stations and graders), and while the reliability of examiner grading was outside the considerations of this project, it may have influenced the results. This project attempted to overcome this issue by randomizing students (and thus their clients and instructors) to the control and intervention groups. Finally, there may also have been contamination between the intervention and control groups, where score reports and information on how to improve were shared between classmates. A next step in this area of research would be to replicate this study with a larger, more diverse group of students.

## CONCLUSION

Providing DSR to students following a dental hygiene SCA resulted in more accurate self-assessments but did not improve performance. Online DSR offers a promising feedback framework and enhancing the reports may facilitate behaviour change. Suggestions include providing links to relevant references, incorporating video feedback, and developing more personalized/individualized reports.

## ACKNOWLEDGEMENTS

This project was supported by CIHR's Frederick Banting and Charles Best Canada Graduate Scholarship, and the University of Alberta's Educational Research and Scholarship Fund.

## CONFLICTS OF INTEREST

The authors declare they have no financial, economic or professional interests that may have influenced the design, execution or presentation of this scholarly work.

APPENDIX A: REFLECTION QUALITY GRADING RUBRIC<sup>a</sup>

Code each comment (each sentence or unique idea)

What? (Thinking about the past) – “What did you do well?”		
Code	Level of reflection	Examples
0	<b>Descriptive:</b> The student simply states what they did without elaboration or analysis of themselves, the client or the context.	“I introduced myself” “I used open-ended questions” “I asked all 10 questions”
1	<b>Analytical:</b> The student interprets the event by adding context: explains why or implications (“so what?”), makes connections to other experiences (e.g., clinic) or coursework/literature.	“I introduced myself to establish a rapport” “I asked questions about her medical condition, so I could make sure there were no contraindications to care”

Now what? (Thinking about the future) – “What could you improve?”		
Code	Level of reflection	Examples
0	<b>Descriptive:</b> The student acknowledges what they need to improve on, but without elaboration on how or why.	“I need to communicate better” “I could speak slower” “Know contraindications to care”
1	<b>Analytical/practical implications:</b> The student demonstrates a deeper understanding by establishing context (“why”) or by describing a specific plan to improve (“how”).	“I need to work on how I communicate with my patients by developing a rapport” “I need to review course materials on all contraindications to care”

<sup>a</sup>Adapted from the University of Alberta HSERC interprofessional reflection guide (2016)

Iterative process of rubric development. Started with HSERC levels of reflection and adapted to suit the style of reflection data.

## APPENDIX B: RUBRIC FOR CODING REFLECTION CONTENT

Domain	Code	Definition	Examples
Communication	C	Anything about improving verbal or non-verbal communication skills.	“Using open-ended questions” “Speaking slower” “Introduce self” “Be more organized”
Client-centred care	CCC	Anything about engaging client in the conversation, or respecting the client’s rights and opinions.	“Respect patient” “Address patient’s chief concern” “Use client’s name”
Eliciting essential information	E	Anything about making sure the proper questions and follow-up questions are used.	“I asked all 10 necessary questions” “Maybe there were more follow-up questions I could have asked”
Interpreting findings	IF	Anything about using the information provided by the client to make care decisions.	“Know contraindications to treatment” “Know when premedication is required”
Other	O	Anything that does not fit into the above categories.	“Prepare better” “I did well overall”

One comment could relate to several different codes. For example, “I need to make sure I am asking all the right questions, so I can make sure there are no contraindications to care” would be coded as both eliciting essential information and interpreting findings.

## REFERENCES

1. Clarke A, Lai H, Sheppard ADE, Yoon MN. Development of diagnostic score reporting for a dental hygiene structured clinical assessment. *Can J Dent Hyg*. 2021;55(1):48–56.
2. Branch WT, Paranjape A. Feedback and reflection: Teaching methods for clinical settings. *Acad Med*. 2002;77(12):1185–1188.
3. Sadler DR. Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*. 2010;35(5):535–50.
4. Ende J. Feedback in clinical medical education. *JAMA*. 1983;250(6):777–81.
5. Van de Ridder JMM, Stokking KM, McGaghie WC, ten Cate OTJ. What is feedback in clinical education? *Med Educ*. 2008;42(2):189–97.
6. Kolb D. The process of experiential learning. In: *Experiential learning: Experience as the source of learning and development*. 2nd ed. Upper Saddle River, New Jersey: Pearson Education; 2015.
7. Torbert WR. The interplay of feedback, attention, and consciousness. In: *Learning from experience: Toward consciousness*. New York, NY: Columbia University Press; 1972.
8. Sandars J. The use of reflection in medical education: AMEE Guide No. 44. *Med Teach*. 2009;31(8):685–95.
9. Archer JC. State of the science in health professional education: effective feedback. *Med Educ*. 2010;44(1):101–108.
10. Clark S, Duggins A. What research guides our beliefs about professional learning? In: *Using quality feedback to guide professional learning: A framework for instructional leaders*. Thousand Oaks, CA: Corwin; 2016.
11. Veloski J, Boex JR, Grasberger MJ, Evans A, Wolfson DB. Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. *Med Teach*. 2006;28(2):117–28.
12. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: Effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. 2012;(6):CD000259.
13. Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol Bull*. 1996;119(2):254–84.
14. Harrison CJ, Molyneux AJ, Blackwell S, Wass VJ. How we give personalised audio feedback after summative OSCEs. *Med Teach*. 2015;37(4):323–26.
15. Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356(4):387–96.
16. Boursicot K, Etheridge L, Setna Z, et al. Performance in assessment: Consensus statement and recommendations from the Ottawa conference. *Med Teach*. 2011;33(5):370–83.
17. Wardman M, Yorke, VC, Hallam JL. Evaluation of a multi-methods approach to the collection and dissemination of feedback on OSCE performance in dental education. *Eur J Dent Educ*. 2018;22(2):e203–e211.
18. Harden R. What is an OSCE? *Med Teach*. 1988;10(1):19–22.
19. Newble D. Techniques for measuring clinical competence: Objective structured clinical examinations. *Med Educ*. 2004;38(2):199–203.
20. Rushforth HE. Objective structured clinical examination (OSCE): Review of literature and implications for nursing education. *Nurse Educ Today*. 2007;27(5):481–90.
21. Blue C. Objective structured clinical exams (OSCE): A basis for evaluating dental hygiene students' interpersonal communication skills. *Access*. 2006;20(7):27–31.
22. Harrison CJ, Könings KD, Molyneux A, Schuwirth LW, Wass V, van der Vleuten CP. Web-based feedback after summative assessment: How do students engage? *Med Educ*. 2013;47(7):734–44.
23. Gotzmann A, De Champlain A, Homayra F, et al. Cheating in OSCEs: The impact of simulated security breaches on OSCE performance. *Teach Learn Med*. 2017;29(1):52–58.
24. Bulut O, Cutumisu M, Aquilina AM, Singh D. Effects of digital score reporting and feedback on students' learning in higher education. *Front Educ*. 2019;4(65).
25. Hollingsworth MA, Richards BF, Frye AW. Description of observer feedback in an objective structured clinical examination and effects on examinees. *Teach Learn Med*. 1994;6(1):49–53.
26. White CB, Ross PT, Gruppen LD. Remediating students' failed OSCE performances at one school: The effects of self-assessment, reflection, and feedback. *Acad Med*. 2009;84(5):651–54.
27. Hendersen M, Ryan, T, Phillips M. The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education*. 2019;44(8):1237–252.
28. Goodman DP, Hambleton RK. Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education*. 2004;17(2):145–220.
29. Roberts MR, Gierl MJ. Developing score reports for cognitive diagnostic assessments. *Educational Measurement: Issues and Practice*. 2010;29(3):25–38.
30. Sinharay S, Puhan G, Haberman SJ. Reporting diagnostic scores in educational testing: Temptations, pitfalls, and some solutions. *Multivariate Behavioral Research*. 2010;45(3):553–73.
31. Zapata-Rivera JD, Katz IR. Keeping your audience in mind: Applying audience analysis to the design of interactive score reports. *Assessment in Education: Principles, Policy & Practice*. 2014;21(4):442–63.
32. Zenisky AL, Hambleton RK. Developing test score reports that work: The process and best practices for effective communication. *Educational Measurement: Issues and Practice*. 2012;31(2):21–26.
33. Roberts MR, Gierl MJ. Development of a framework for diagnostic score reporting. Paper presented at the Annual meeting of the American Educational Research Association, San Diego, CA 2009.
34. Gotch C, Roberts RR. A review of recent research on individual-level score reports. *Educational Measurement: Issues and Practice*. 2018;37(3):46–54.
35. Pickett FA. Personal, dental, and health histories. In: Darby ML, Walsh MM, eds. *Dental hygiene theory and practice*. 3rd ed. St. Louis, MO: Saunders Elsevier; 2010:149–79.
36. Canadian Dental Hygienists Association (CDHA), Federation of Dental Hygiene Regulatory Authorities (FDHRA), Commission on Dental Accreditation of Canada (CDAC), National Dental Hygiene Certification Board (NDHCB), Dental Hygiene Educators of Canada (DHEC). *Entry-to-practice competencies and standards for Canadian dental hygienists*. Ottawa, ON: CDHA; 2010.
37. Darby ML, Walsh MM. *Dental hygiene theory and practice*. 3rd ed. St. Louis, MO: Saunders Elsevier; 2010.



38. Hsu C-C, Sandford BA. The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*. 2007;12(10):1–8.
39. StataCorp. Stata Statistical Software: Release 14. In. College Station, TX: StataCorp; 2015.
40. Senn S. Change from baseline and analysis of covariance revisited. *Stat Med*. 2006;25(24):4334–344.
41. Health Sciences Education and Research Commons. Interprofessional Reflection Guide. 2016; Available from: <http://www.hserc.ualberta.ca/Resources/CurricularResources/InterprofessionalReflectionGuide.aspx>
42. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74.
43. Hanson EC. Analysing qualitative data. In: *Successful qualitative health research: A practical introduction*. New York, NY: Open University Press; 2006:137–60.
44. Webster M, Remedios L. Reflection and feedback during my OSCE: What have I learnt? *Physiotherapy*. 2015;101:e1273.
45. Quinton S, Smallbone T. Feeding forward: Using feedback to promote student reflection and learning—a teaching model. *Innovations in Education and Teaching International*. 2010;47(1):125–35.
46. Zimmerman BJ. Self-regulated learning and academic achievement: An overview. *Educational Psychologist*. 1990;25(1):3–17.
47. Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*. 2006;31(2):199–218.
48. Butler DL, Winne PH. Feedback and self-regulated learning: A theoretical synthesis. *Rev Educ Res*. 1995(3):245.
49. Trout DL, Hyde E. Developing score reports for statewide assessments that are valued and used: Feedback from K-12 stakeholders. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA 2006.
50. Fukkink RG, Trienekens N, Kramer LJ. Video feedback in education and training: Putting learning in the picture. *Educ Psychol Rev*. 2011;23(1):45–63.
51. Massey D, Byrne J, Higgins N, Weeks B, Shuker M-A, Coyne E, et al. Enhancing OSCE preparedness with video exemplars in undergraduate nursing students. A mixed method study. *Nurse Educ Today*. 2017;54.