

Multiple feedback sources in learning clinical history-taking skills: Developing evaluative judgement

K. Fitzgerald^{1,2}, B. Vaughan¹ & J. H. M. Tai³

Abstract

Introduction: Contemporary feedback models emphasise the value of multiple feedback opportunities. Effective feedback participation requires evaluative judgement—the ability to discern the quality of one’s own and others’ work. Self and peer assessment may enable repeated practice and feedback for developing evaluative judgement. However, attitudes to self and peer assessment may present a barrier to effective implementation. This study explored whether congruence between marks from self and peer assessment improved with assessment task participation. Participants’ attitudes towards self and peer assessment and approaches to learning were also evaluated.

Methods: Participants undertook simulated history-taking tasks in semester 2, 2018. Group 1 undertook formative and summative assessments and participated in self and peer assessment. Group 2 undertook formative and summative assessment. Group 3 undertook only summative assessment. All groups received faculty feedback for each submitted assessment. Participants completed the modified Study Process Questionnaire (mSPQ) and the Peer Perception of Assessment (PPA) before (T1) and after the formative task (T2) and after the summative task (T3).

Results: Summative task scores improved for group 1 ($n = 9$, $p < 0.01$) and group 2 ($n = 26$, $p < 0.01$). Within-group ($p = 0.02$) and between-group differences ($p = 0.01$) were identified for surface learning approaches. All groups’ perceptions of peer assessment decreased significantly ($p < 0.01$) across all three time periods.

¹ Department of Medical Education, University of Melbourne, Melbourne, Victoria

² School of Health and Biomedical Sciences, RMIT University, Melbourne, Victoria

³ Centre for Research in Assessment and Digital Learning, Deakin University, Melbourne, Victoria

Correspondence

Kylie Fitzgerald
Department of Medical Education
The University of Melbourne
Parkville, VIC 3010
Australia
Ph: +61 4 6692 2032
Email: kylie.fitzgerald@unimelb.edu.au

Conclusions: Participants receiving self and peer assessment and faculty feedback improved performance and increased congruence of their self- and peer-assessment marks, potentially developing their evaluative judgement skills. Peer assessment perception became less positive, while surface learning approaches increased. Future research should assess the role of self and peer assessment in developing evaluative judgement.

Keywords: evaluative judgement; self-assessment; feedback; medical history taking

Introduction

Clinical skills are an essential component of health professional education and practice. One challenge of teaching clinical skills is creating productive learning environments with adequate opportunities and resourcing to support engagement in meaningful feedback. Opportunities for feedback to improve practice are limited, especially in classroom-based teaching, where large cohort sizes may limit or prevent this. The need for multiple feedback sources, including faculty, self and peer, has been investigated in medical education (Anwer et al., 2017; Beaudoin et al., 2019) and warrants further investigation in additional contexts.

Though peer feedback has been found to be valuable for learning (Adachi et al., 2017; Rush et al., 2012; Tai, Canny, et al., 2017), concerns have been expressed regarding the accuracy of self- and same-level peer assessments (Lanning et al., 2011; Mehrdad et al., 2012). In particular, where learners have little prior experience with the tasks and limited understanding of the area in which they are to provide/receive feedback, the credibility of these judgements may be questioned (Ahmed et al., 2013). Synthesising multiple sources of feedback may help to address these concerns with respect to consistency and fairness (Tai, Canny, et al., 2016). Multiple sources may also assist in improving learners' understanding of quality work, improve self-assessment calibration and provide feedback on different aspects of task performance.

This strategy draws on the "Feedback Mark 2" model proposed by Molloy & Boud (2013), where learners have a number of opportunities to improve the quality of their work through feedback.

Feedback is also essential for the development of learners' evaluative judgement (Carless & Boud, 2018). Evaluative judgement is defined as the capability of the learner to assess the work of self and others and involves both the understanding of quality and the application of that understanding in judgements of work (Tai, Canny, et al., 2016). It may be evident through an individual's judgement of a piece of work. However, the development of evaluative judgement may also be implied where an individual improves the quality of their own work over time, having cultivated a better understanding of the quality requirements for that task.

Pedagogical practices that may develop evaluative judgement include the use of scaffolds, such as rubrics or exemplars, to establish a shared understanding of quality work. Self and

peer assessment has been suggested as a means for learners to practise making judgements. However, it is not known if multiple opportunities for self and peer assessment lead to more accurate and nuanced assessments, i.e., improvements in learners' evaluative judgement.

Focusing on self and peer assessment may be beneficial, since it provides opportunities to engage in feedback processes and may facilitate development of learners' evaluative judgement (Tai, Ajjawi, et al., 2018; Tai, Canny, et al., 2016). Learner attitudes towards self and peer assessment may present a barrier to implementation, however positive attitudes can also be developed through training and experience in such assessment tasks (van Zundert et al., 2010). Since benefits from peer assessment accrue for both parties, learners should be able to benefit from participation (Yu et al., 2011). Previous investigation of attitudes towards peer assessment in an osteopathy student cohort demonstrated mixed views towards the use of peer marking (Fitzgerald & Vaughan, 2018), and this warrants further investigation.

Students' engagement in self and peer assessment may also be influenced by the way they approach learning in general, given it requires significant involvement in learning activities. Biggs et al. (2001) identified three approaches to learning: deep, surface, and strategic. Surface approaches minimise effort and learning while satisfying basic assessment requirements, while deep approaches are appropriate and intrinsic to learning and strategic (originally referred to as achieving) approaches are focused on short-term outcomes (Biggs, 1993). Previous studies of approaches to learning within health professional education have focused on cross-sectional correlations with written examinations and study techniques rather than learning of clinical skills (Delgado et al., 2018; Kamath et al., 2018; Malik et al., 2019; Svirko & Mellanby, 2008). Within clinical skills training, a deep approach may be important both for learning and for participation in self and peer assessment, since learners should consider both the requirements of their present task and developing skills that will be consistent with the requirements of their future health professional practice.

While self and peer assessment may support the development of evaluative judgement, further investigation is needed given the relative lack of published evidence to date. We also should consider the impact of learners' motivations to engage in deep learning and their attitudes towards learning. Therefore, this study aimed to explore the role of self and peer assessment in the development of evaluative judgement, in addition to faculty feedback, within clinical skills teaching assessments. The three research questions were:

1. Does the congruence between marks from self and peer assessment (suggesting increased evaluative judgement capacity) improve with practice?
2. How do the participants' perceptions of self and peer assessment change through their engagement in these approaches?

3. Is there a relationship between participants' approaches to learning and their participation in self and peer assessment?

Methods

A prospective quantitative study was conducted involving self, peer and faculty assessments of formative and summative assessment tasks. Potential participants were informed of the study via face-to-face information sessions and emails that were sent to the entire cohort (Figure 1). This study was approved by the Victoria University Human Research Ethics Committee (HR17-178).

Participants

All full-time year 2 participants enrolled in the unit Clinical Skills 4 had the opportunity to participate. Learners were able to self-select into the three groups by choosing their level of participation or by not participating at all.

- Group 1 participated in both the formative and summative assessment tasks and undertook self and peer marking and received faculty feedback for both tasks (Figure 1).
- Group 2 participated in the formative and summative assessment tasks and received faculty feedback for both tasks but did not undertake self or peer marking for either task.
- Group 3 participated in the summative assessment task only and received faculty feedback for it.

Therefore, group allocation was based on completed assessment task and marking submissions. For example, only participants who had submitted their self and peer assessment for both the formative and summative assessment tasks were included in group 1 analyses.

The study is described in two parts: Part 1—assessment task and Part 2—evaluation of self- and peer-assessment perceptions and participant self-report study process.

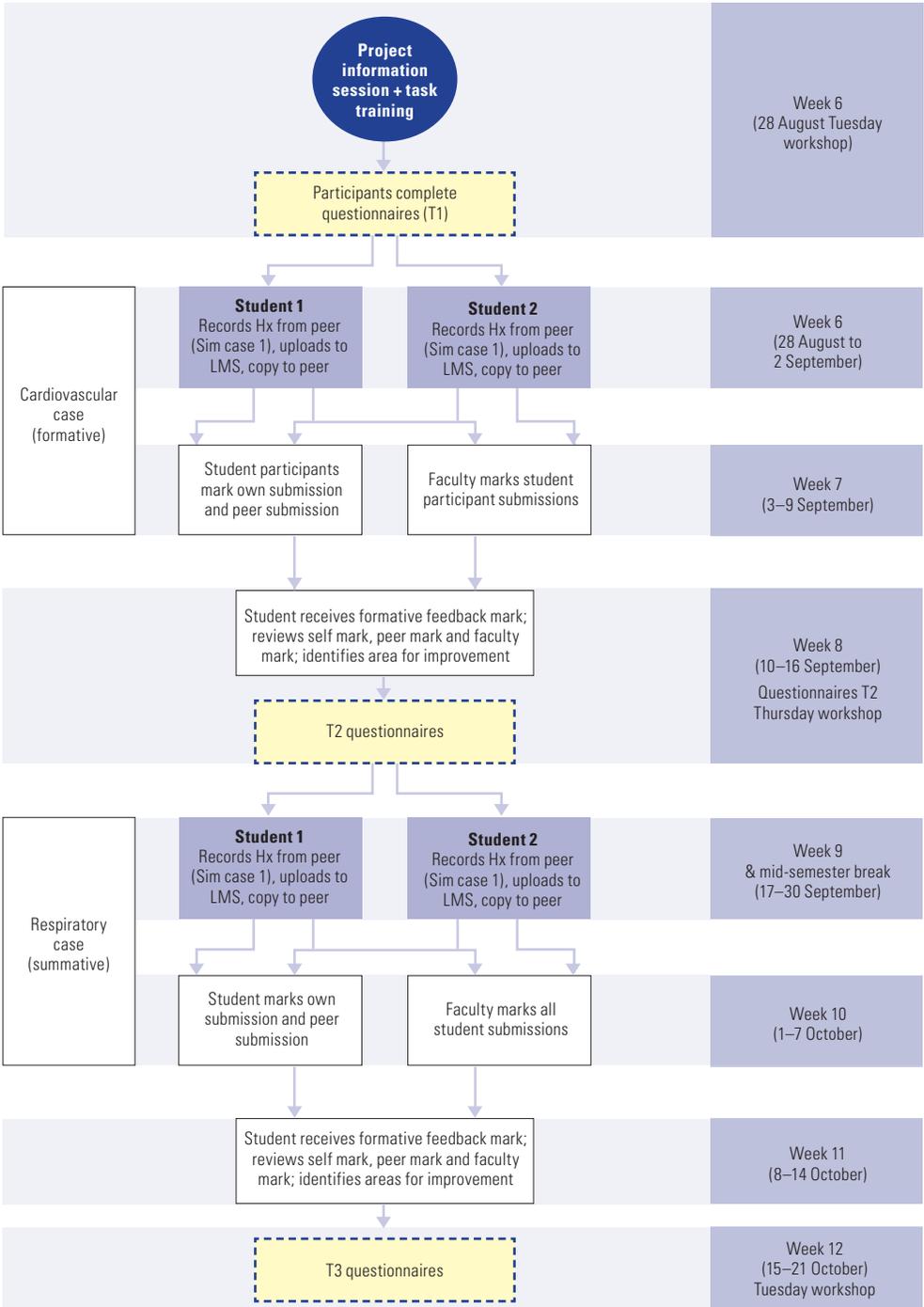
Part 1:

Participation in formative assessment prior to submitting summative assessment

Group 1 and 2 participants undertook a simulated history-taking assessment task twice, with an option for multisource feedback between tasks. The assessment was undertaken in Year 2 of the Bachelor of Science (Osteopathy) at Victoria University (Melbourne, Australia) in semester 2, 2018. The study design outlined below was informed by the key elements of the Feedback Mark 2 model (Boud & Molloy, 2013) to create an approach that is agentic, featuring an adaptive/responsive system where the effects are monitored by teachers and learners and where there is a double feedback loop with the goal of judgement performance improvement.

Figure 1

Group 1 Participant Process



Educational intervention: Assessment task 1 (Formative)

Group 1 and 2 participants worked in pairs. Participant 1 acted as the practitioner and completed the history-taking component of an osteopathic consultation using a simulated case scenario. Participant 2 acted as the simulated patient. All participants had previously received extensive training in simulated clinical scenarios as practitioners and simulated patients and had completed similar simulation tasks.

Participant 1 used an electronic device to record a video of up to 10 minutes in length of the simulated clinical interaction. The simulated clinical interaction was undertaken at a time of their choosing in a variety of settings, including homes or in the Victoria University Osteopathy Clinic (the on-campus, student-led clinic). Participants then swapped roles and completed a different simulated clinical scenario. Both scenarios were based on cardiovascular medicine, which aligned with the curriculum content during those weeks of the semester.

Using the SHARP debriefing tool (Arora et al., 2011), Each participant reviewed the video recording of them acting as the practitioner (Table 1) and submitted short responses using the SHARP prompts. This tool directs the participant to reflect upon and critically analyse their own performance via a directed acronym.

Table 1

SHARP Feedback Tool

SHARP Acronym	
S	Set learning objectives for task.
H	How did it go? What went well and why?
A	Address concerns. What did not go so well and why?
R	Review learning points. Were your learning objectives met for this case?
P	Plan ahead. What actions can you take to improve your future practice?

Participants then uploaded their video and written reflection for faculty marking. Group 1 participants also shared their video and reflections with their peer participant for the purpose of peer marking.

All participants were provided with an in-class training session to explain the rubric structure (used for both formative and summative tasks) and marking. Faculty assessors received a video module containing task information, which demonstrated the use of rubrics to train them in the task. Two example submissions in the module showed a pass grade submission (55%) and the other a distinction (75% mark) submission. Group 1 participants completed self-marking of their own submission and marked their partner’s submission using the same rubric.

Within 2 weeks, each participant received their faculty-marked formative assessment rubric with feedback on their submission. Group 1 participants also received their self- and peer-marked rubric and feedback. Participants were prompted to review and reflect on their feedback within 2 weeks and utilise this feedback to develop areas of their performance that were identified as needing improvement from their multisource feedback (faculty, self, peer assessment).

Educational intervention: Assessment task 2 (Summative)

The process for the summative assessment task for those in group 1 and 2 was identical to the formative assessment task described above. The simulated scenarios for this iteration were based upon respiratory health, in alignment with the curriculum content during those weeks of the semester. Participants who did not participate in the self- and peer-assessment component of the study (group 3) completed the summative assessment only. This component is the same as described in the formative assessment except that these participants did not self or peer mark their submission. Group 3 participants received only summative faculty marks and feedback. They had access to the cardiovascular scenarios used in the formative task but, in choosing to be in group 3, had elected not to undertake deliberate practice and reflection by completing the formative task and receiving feedback prior to submitting their summative assessment.

Rubric design

Congruence of marks was taken as a proxy for the development of evaluative judgement. To compare marks across assessors, all assessors (self, peer and faculty) used the same rubric to both grade the task and provide feedback. Marking criteria aligned with each of the five steps used in the SHARP feedback tool (Table 1). For all participants who submitted formative and/or summative assessments, the total marks of all assessors were added together. This final mark was used to assess differences between the groups.

Part 2:

Measures of participant perceptions of peer assessment and approaches to learning

Participants

All full-time year 2 participants enrolled in the unit Clinical Skills 4 had the opportunity to participate in Part 2 of this project by completing questionnaires handed out at T1 (before formative assessment, Task 1), T2 (after formative assessment, Task 1 and feedback received) and T3 (after summative assessment, Task 2 and feedback received) (see Figure 1).

To assess how the participants' perceptions of self and peer assessment changed through their engagement in these approaches, the Participants' Perceptions of and Attitudes Towards Peer Assessment (PPA) survey was used. The PPA is a 20-item questionnaire, developed by Wen and Tsai (2006), that assesses points of view towards and experiences

of participating in peer assessment. Some items were removed from the original questionnaire as they referenced online learning and were not relevant to this study.

To identify if there was a relationship between participants' approaches to learning and their participation in self and peer assessment, the Study Process Questionnaire (SPQ) was used. SPQ reliability measures (Cronbach α) of the scales and subscales have been reported as ranging from the low 50s to high 70s (Snelgrove & Slater, 2003). Since the SPQ measures general approaches to learning (Biggs, 1993; Biggs et al., 2001), it was adapted to the osteopathy participant context (mSPQ), and this adaptation has undergone confirmatory factor analysis in osteopathy student populations (Vaughan, 2016).

Data analysis

Data were extracted from each of the assessment rubrics and questionnaires. Descriptive statistics were generated for the assessments and completed questionnaires. Inferential statistics were used to explore differences for participation groups (chi-square), between marker groups and pre-post (non-parametric) assessment marks, and within and between-group differences for the mSPQ and PPA (repeated measures ANOVA). Alpha was set at $p < 0.05$ and effect sizes (Cohen's d and partial eta-squared for parametric inferential statistics, r , for non-parametric) calculated where relevant. Pearson's r was used to evaluate the relationship between the mSPQ subscales and PPA.

Results

One hundred and thirteen participants were enrolled in the clinical skills subject, and 56 participants participated in the clinical history-taking task, with written consent provided for access to their marks for these tasks. Group 1 had nine participants; group 2 had 26 participants; and group 3 had 21 participants.

Assessment results

Most participants were female ($n = 39$, 69%), with 24 participants (40.7%) aged 18–20 years and 18 (30.5%) aged 21–23 years. No significant difference ($p > 0.05$) was identified for gender or age with respect to participation in any group. Descriptive statistics for the assessment marks for both groups are presented in Table 2.

Within group comparisons demonstrated that group 1 (intervention) formative marks showed no significant difference between faculty mark and self-mark ($p > 0.05$).

Significant differences were identified between the self-mark and peer mark ($p = 0.01$, $z = -2.50$, $r = 0.83$) and the faculty mark and peer mark ($p = 0.04$, $z = -1.97$, $r = 0.65$), with the peer mark being higher in both combinations. Group 1's faculty mark increased significantly from their formative to their summative mark (group 1: $p < 0.01$, $z = -2.40$, $r = 0.80$), as did group 2 ($p < 0.01$, $z = -3.6$, $r = 0.71$), with large effect sizes. Between-group comparisons revealed no significant difference between groups 1 and 2 for the formative

assessment task faculty mark. For the summative assessment task faculty mark, group 1 and 2 showed improved and higher marks than group 3, though the difference was not statistically significant.

Table 2

Mean Assessment Task Marks Across the Three Marker Groups

Marker Group	Assessment Component	Faculty Mark	Self Mark	Peer Mark
		Total score out of 30 marks Mean (+/- SD)		
Group 1 (n = 9)	Formative	18.3 (+/- 1.6) ^a	18.1 (+/- 2.0) ^b	20.5 (+/- 3.0)
	Summative	21.7 (+/- 2.3) ^c	22.5 (+/- 2.8)	23.7 (+/- 3.0)
Group 2 (n = 26)	Formative	15.7 (+/- 4.8)	-	
	Summative	20.2 (+/- 2.2) ^c	-	
Group 3 (n = 21)	Summative	18.9 (+/- 2.9)	-	

Notes:

^a faculty versus peer, $p = 0.04$

^b self versus peer mark, $p = 0.01$

^c faculty formative versus summative, $p < 0.01$

Modified Study Process Questionnaire (mSPQ)

Matched data were available at T1, T2 and T3 for the 59 participants who completed the mSPQ and the PPA (52% of cohort). Thirty five (62.5%) of those completing the questionnaires indicated they had consented to providing access to their marks (thus were in group 1, 2 or 3). Most participants in these groups were female (n = 24; 68.6%), with 13 participants (37.1%) aged 18–20 years and 8 (22.9%) aged 21–23 years. No significant difference was identified for gender or age between the participation group and the non-participation group.

Descriptive statistics are presented in Table 3. For the deep learning subscale, no longitudinal within-group difference was demonstrated. A significant between-group difference was demonstrated for the participant group reporting higher mean scores for the deep learning subscale across all three time periods ($p = 0.03$, $F = 4.84$, $eta = 0.08$). For the surface learning subscale, significant within-group ($p = 0.02$, $F = 6.08$, $eta = 0.1$) and between-group differences ($p = 0.01$, $F = 6.35$, $eta = 0.1$) were identified, with the participants’ group reporting lower mean scores. Plots for the mSPQ deep and surface learning subscale are presented in Figures 2 and 3, respectively.

Table 3

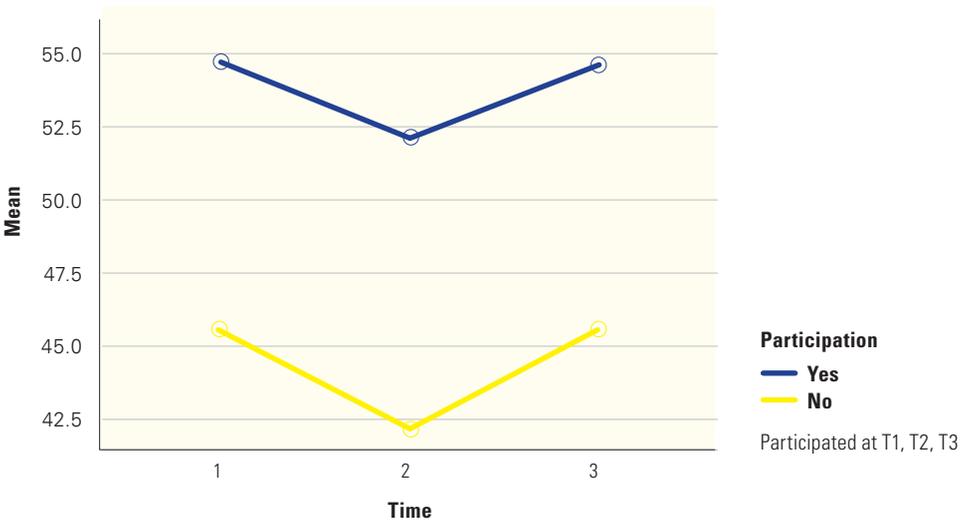
Descriptive Statistics for the Deep and Surface Learning Subscales of the mSPQ

Participating	Deep Learning Subscale		Surface Learning Subscale	
	Yes (n = 35)	No (n = 24)	Yes (n = 35)	No (n = 24)
Time 1	54.8 (19.4)	45.5 (15.1)	17.5 (10.1)	23.4 (14.2)
Time 2	52.1 (20.4)	42.2 (17.7)	28.0 (17.7)	45.2 (22.8)
Time 3	54.6 (19.2)	45.5 (17.4)	37.2 (21.4)	43.6 (20.0)

Note: Values represent an ordinal to interval conversion score described by Vaughan (2016)

Figure 2

Mean M-SPQ Deep Subscale Scores Over the Period of the Study



Perception of peer assessment

Participants’ perceptions of peer assessment decreased significantly within groups ($p < 0.01$, $F = 40.28$, $\eta^2 = 0.29$) across all time periods with no between-group difference (Figure 4). Participating in multisource feedback (group 1) was not associated with previous experience with peer assessment ($p > 0.05$) nor was the percentage of the grade attributed to the peer assessment ($p > 0.05$). Perception of peer assessment was minimally associated with mSPQ subscales at their respective time periods ($r < 0.30$). No correlation was statistically significant.

Figure 3

Mean M-SPQ Surface Subscale Scores Over the Period of the Study

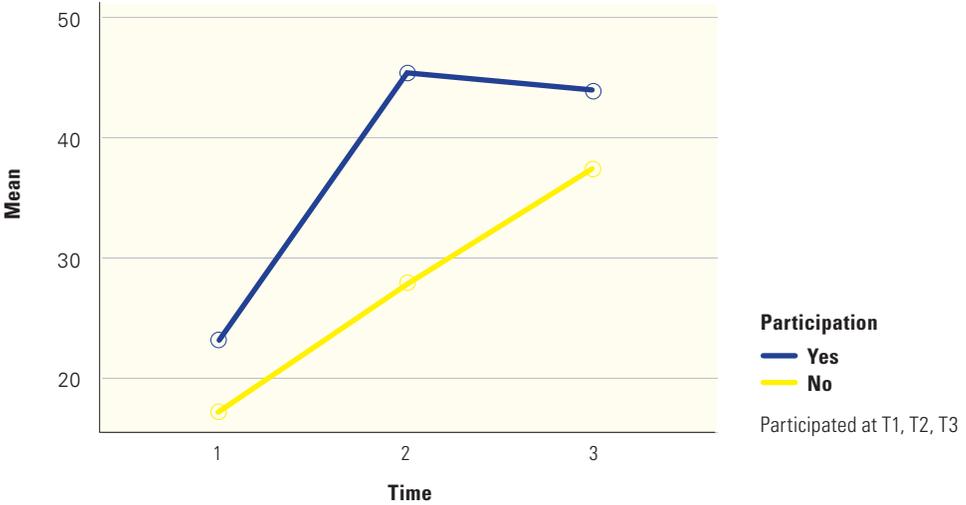
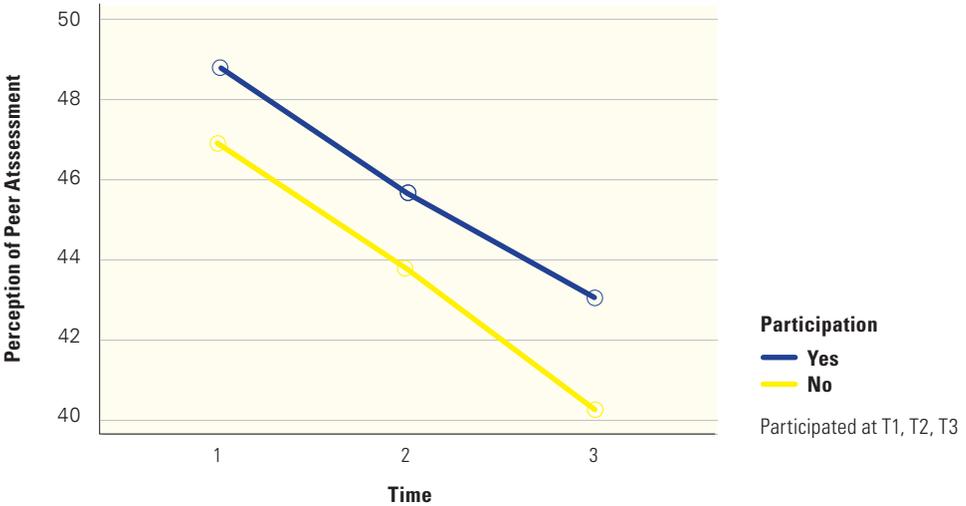


Figure 4

Participants' Perceptions of Peer Assessment Over the Period of the Study



Discussion

Our findings enable exploration of the role of self and peer assessment in developing evaluative judgement in a simulated clinical skills history-taking assessment task. Participants who undertook self and peer assessment improved their marks between the

formative and summative assessments, as did those who only received faculty feedback between the formative and summative assessments. Perceptions of peer assessment became less positive over the course of this study for all participants, with surface learning approaches also becoming more pronounced over the study.

Development of evaluative judgement

In this study, we consider if evaluative judgement capacity development might be demonstrated by improved congruence of group 1 participants' self, peer and faculty grades, i.e., the development of shared understanding of quality, where quality can be represented through a rubric. The opportunity to practise assessment of self and others has been suggested to be important, since it transforms the activity from grade guessing to a meaningful learning activity with the goal of improving task performance and evaluative judgement (Tai, Canny, et al., 2017).

Peer-assisted learning has been identified by learners in medical clinical placements as contributing to their understanding of concepts of quality and undertaking comparisons between an observed performance and the assessment standards (Tai, Canny, et al., 2016). In this study, the design of the learning activity, rubric and scaffolding of reflection through the SHARP tool could have offered learners methods for establishing shared understandings of the task standards.

Participants in group 2 (only faculty formative and summative assessment feedback) also improved their mark for the assessment task. This aligns well with previous research that formative assessment feedback from a reputable source is helpful for learning and subsequent task performance (Hattie & Timperley, 2007). Although group 2 did not explicitly undertake self-assessment using the rubric, it is possible the faculty feedback stimulated self-assessment. This may have guided the participants to improve their performance on the subsequent task, also improving their evaluative judgement.

This is consistent with Feedback Mark 2 strategies about comparative judgements and planning for improvements in future task performance (Boud & Molloy, 2013). Why students chose to participate in their respective groups could be related to perceptions of the credibility of the feedback source(s). How we can go about improving students' understanding of the importance of participating in feedback and judging the credibility of feedback sources is an area that requires further research (Molloy et al., 2020).

Approaches to learning

The mSPQ results suggest that those who completed this questionnaire and who reported a deeper learning strategy were more likely to participate in group 1 (voluntary self- and peer-assessment activities). The trend of increasing surface scores for group 1 participants might indicate a switch in strategies in response to increasing course and assessment demands across the semester, i.e., the influence of the context of learning as opposed to fixed participant characteristics (Biggs et al., 2001).

Our findings contrast with recent studies reporting a decrease in surface learning approaches (Falk et al., 2016) and an increase in deep and strategic learning approaches over time (McDonald et al., 2017). However, the timescale of these studies was much longer than our research, which was conducted within one semester. Future work could longitudinally investigate approaches to learning in other settings and connections between motivation to participate in peer-assisted learning and study processes.

Participants' perceptions of peer assessment

All participants' perceptions of peer assessment scores decreased over time, indicating a worsening perception of peer assessment. One possibility for the shift might be the need to focus on surface learning strategies for assessments when under pressure rather than taking a deeper and broader approach to learning. Peer assessment may not be seen as a fruitful learning strategy when participants are under pressure to learn materials for examinations and submit other final assessment tasks. It is likely there are other variables that would influence perception that are beyond the scope of this project, and qualitative approaches would be useful to explore the change in perceptions and reasons for it further.

The finding of negative perceptions towards peer assessment aligns with previous findings in osteopathic education (Fitzgerald & Vaughan, 2018). This suggests that curricula-embedded efforts to socialise the participants into discipline practices of self and peer evaluation are insufficient. Future work could explore opportunities for specific and explicit education on self and peer assessment within the curriculum.

Reflections on feasibility

The group 1 participants who participated in peer- and self-assessment tasks improved alignment of their judgement with the faculty marking. While there was an improvement in their score on the summative task compared to students who did not participate in self or peer assessment, this was not statistically significant. This may be due to the small sample size of group 1 ($n = 9$). However, it may alternatively suggest that peer- and self-assessment practice, while helpful for students to meet course "norms", does not alone result in exceptional performance. Consideration should be given to the feasibility of this approach with respect to student workload and improvement in performance. The time and effort required by the faculty also needs to be considered.

Limitations

A limitation of this study is the small and unequal sample sizes. The pragmatic nature of the study design allowed participants to opt into one of the three groups, which led to unequal group sizes. Furthermore, we were unable to determine if there were prior differences between groups in their clinical history-taking abilities. A randomised study design with equal group sizes may be useful in future work. The intervention was also of a limited duration, with no follow-up. Additional opportunities to practise self and peer assessment with a larger participant cohort or across different learner levels would

enable further exploration of the role of different sources of feedback to develop evaluative judgement.

This research used previously validated instruments (mSPQ, PPA) to investigate participants' study processes and perceptions of peer learning. The PPA has not been validated within an osteopathy participant population and was modified as the original measures contained irrelevant items (e.g., perceptions of online learning). These amendments to the PPA may limit the validity of the questionnaire. Further work may be required to establish the measurement properties of the questionnaires (i.e., retest reliability) as well as further explore the relationship of deep and surface learning to evaluative judgement. Future studies may wish to develop new measurement instruments or draw upon qualitative approaches to explore this further.

The opportunity to receive concurrent faculty feedback may also have influenced participants' perceptions. Our use of faculty marks as the "gold standard" against which we compared participant marks may ignore the useful contributions on other aspects of performance that participants can make within assessments. Though we believe the rubric used required holistic judgements of quality, the site of action of self and peer assessment in formative assessment feedback may extend beyond the boundaries of the assessment tools we provided.

Conclusion

Participants who received multiple sources of feedback demonstrated improved performance when repeating a clinical history-taking assessment. These participants also demonstrated increased congruence of their self and peer marks with faculty marks, potentially developing their evaluative judgement skills. Improvement in overall performance was also observed for participants who received only faculty feedback for the formative task, so further exploration of the value of self and peer assessment in this context is needed. Perceptions of peer assessment became less positive, while surface learning approaches were more pronounced throughout the study. Our findings suggest that participants might perceive formative assessment tasks and/or self and peer assessment as burdensome, rather than as a learning opportunity, especially closer to end-of-semester assessment. Further evaluation of the inclusion of self and peer assessment is warranted.

Conflicts of interest and funding

The authors declare no conflicts of interest or funding for this project.

References

- Adachi, C., Tai, J. H.-M., & Dawson, P. (2017). Academics' perceptions of the benefits and challenges of self and peer assessment in higher education. *Assessment & Evaluation in Higher Education*, 43(2), 294–306. <https://doi.org/10.1080/02602938.2017.1339775>
- Ahmed, M., Arora, S., Russ, S., Darzi, A., Vincent, C., & Sevdalis, N. (2013). Operation debrief: A SHARP improvement in performance feedback in the operating room. *Annals of Surgery*, 258(6), 958–963. <https://doi.org/10.1097/SLA.0b013e31828c88fc>
- Anwer, L. A., Shareef, M. A., Nurhussen, A., Alkabbani, H. M., Alzahrani, A. A., Obad, A. S., Zafar, M., & Afsar, N. A. (2017). Objectivity in subjectivity: Do students' self and peer assessments correlate with examiners' subjective and objective assessment in clinical skills? A prospective study. *BMJ Open*, 7(5), e012289.
- Arora, S., Runnacles, J., Ahmed, M., Sevdalis, N., Nestel, D., Paige, N., Hull, L., Russ, S., Wheelock, S., Miskovic, D., Darzi, A., & Vincent, C. (2011). *The London handbook for debriefing*. Imperial College London. <https://docplayer.net/21017622-The-london-handbook-for-debriefing-in-clinical-and-simulated-settings-supported-by.html>
- Beaudoin, P.-L., Labbé, M., Fanous, A., Young, M., Rappaport, J., Park, Y. S., Manoukian, J., & Nguyen, L. H. (2019). Teaching communication skills to OTL-HNS residents: Multisource feedback and simulated scenarios. *Journal of Otolaryngology—Head & Neck Surgery*, 48(1), 1–7. <https://doi.org/10.1186/s40463-019-0329-8>
- Biggs, J., Kember, D., & Leung, D. Y. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71(1), 133–149. <https://doi.org/10.1348/000709901158433>
- Biggs, J. B. (1993). From theory to practice: A cognitive systems approach. *Higher Education Research and Development*, 12(1), 73–85. <https://doi.org/10.1080/0729436930120107>
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698–712. <https://doi.org/10.1080/02602938.2012.691462>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>

- Delgado, Á. H. d. A., Almeida, J. P. R., Mendes, L. S. B., Oliveira, I. N. d., Ezequiel, O. d. S., Lucchetti, A. L. G., & Lucchetti, G. (2018). Are surface and deep learning approaches associated with study patterns and choices among medical students? A cross-sectional study. *Sao Paulo Medical Journal*, 136(5), 414–420. <https://doi.org/10.1590/1516-3180.2018.0200060818>
- Falk, K., Falk, H., & Ung, E. J. (2016). When practice precedes theory: A mixed methods evaluation of students' learning experiences in an undergraduate study program in nursing. *Nurse Education in Practice*, 16(1), 14–19. <https://doi.org/10.1016/j.nepr.2015.05.010>
- Fitzgerald, K., & Vaughan, B. (2018). Learning through multiple lenses: Analysis of self, peer, nearpeer, and faculty assessments of a clinical history-taking task in Australia. *Journal of Educational Evaluation for Health Professions*, 15, Article 22. <https://doi.org/10.3352/jeehp.2018.15.22>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Kamath, A., Rao, R., Shenoy, P. J., & Ullal, S. D. (2018). Approaches to learning and academic performance in pharmacology among second-year undergraduate medical students. *Scientia Medica*, 28(4), 32395. <https://doi.org/10.15448/1980-6108.2018.4.32395>
- Lanning, S. K., Brickhouse, T. H., Gunsolley, J. C., Ranson, S. L., & Willett, R. M. (2011). Communication skills instruction: An analysis of self, peer-group, student instructors and faculty assessment. *Patient Education and Counseling*, 83(2), 145–151. <https://doi.org/10.1016/j.pec.2010.06.024>
- Malik, A. A., Khan, R. A., Malik, H. N., Humayun, A., Butt, N. S., & Baig, M. (2019). Assessing reliability and validity of Revised Biggs Two-Factor Study Process Questionnaire to measure learning approaches among undergraduate medical students in Lahore, Pakistan. *The Journal of the Pakistan Medical Association*, 69(3), 337–342.
- McDonald, F., Reynolds, J., Bixley, A., & Spronken-Smith, R. (2017). Changes in approaches to learning over three years of university undergraduate study. *Teaching & Learning Inquiry*, 5(2), 65–79. <https://doi.org/10.20343/teachlearninqu.5.2.6>
- Mehrdad, N., Bigdeli, S., & Ebrahimi, H. (2012). A comparative study on self, peer and teacher evaluation to evaluate clinical skills of nursing students. *Procedia—Social and Behavioral Sciences*, 47, 1847–1852. <https://doi.org/10.1016/j.sbspro.2012.06.911>

- Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*, 45(4), 527–540. <https://doi.org/10.1080/02602938.2019.1667955>
- Rush, S., Firth, T., Burke, L., & Marks-Maran, D. (2012). Implementation and evaluation of peer assessment of clinical skills for first year student nurses. *Nurse Education in Practice*, 12(4), 219–226. <https://doi.org/10.1016/j.nepr.2012.01.014>
- Snelgrove, S., & Slater, J. (2003). Approaches to learning: Psychometric testing of a study process questionnaire. *Journal of Advanced Nursing*, 43(5), 496–505. <https://doi.org/10.1046/j.1365-2648.2003.02747.x>
- Svirko, E., & Mellanby, J. (2008). Attitudes to e-learning, learning style and achievement in learning neuroanatomy by medical students. *Medical Teacher*, 30(9–10), e219–e227. <https://doi.org/10.1080/01421590802334275>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481. <https://doi.org/10.1007/s10734-017-0220-3>
- Tai, J. H., Canny, B. J., Haines, T. P., & Molloy, E. K. (2016). The role of peer-assisted learning in building evaluative judgement: Opportunities in clinical medical education. *Advances in Health Sciences Education*, 21(3), 659–676. <https://doi.org/10.1007/s10459-015-9659-0>
- Tai, J. H., Canny, B. J., Haines, T. P., & Molloy, E. K. (2017). Implementing peer learning in clinical education: A framework to address challenges in the “real world”. *Teaching and Learning in Medicine*, 29(2), 162–172. <https://doi.org/10.1080/10401334.2016.1247000>
- van Zundert, M., Sluijsmans, D., & van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20(4), 270–279. <https://doi.org/10.1016/j.learninstruc.2009.08.004>
- Vaughan, B. (2016). Confirmatory factor analysis of the Study Process Questionnaire in an Australian osteopathy student population. *International Journal of Osteopathic Medicine*, 20, 62–67. <https://doi.org/10.1016/j.ijosm.2016.03.001>
- Wen, M. L., & Tsai, C.-C. (2006). University students’ perceptions of and attitudes toward (online) peer assessment. *Higher Education*, 51(1), 27–44. <https://doi.org/10.1007/s10734-004-6375-8>

Yu, T.-C., Wilson, N. C., Singh, P. P., Lemanu, D. P., Hawken, S. J., & Hill, A. G. (2011). Medical students-as-teachers: A systematic review of peer-assisted teaching during medical school. *Advances in Medical Education and Practice, 2*, 157–172. <https://doi.org/10.2147/AMEP.S14383>