Peer assessment and professional behaviours: what should we be assessing, how, and why?

This is the accepted manuscript.

This is the peer reviewed version of the following article:


which has been published in final form at http://www.dx.doi.org/10.1111/medu.13254.

© 2017, John Wiley & Sons and The Association for the Study of Medical Education

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Downloaded from DRO: http://hdl.handle.net/10536/DRO/DU:30094555
Peer assessment and professional behaviours: what should we be assessing, how, and why?

Joanna Tai & Chie Adachi

Professional behaviour has been long argued as being better assessed by peers than tutors. While many studies published on peer assessment have focussed on professional behaviour, concerns about reliability and validity of student marking remain as one of the widely discussed challenges (1,2). [authors]’ work rightfully challenges the assumptions many of us have previously made about peer assessments of professional behaviour within medical education, namely that it can be a reliable and valid measurement for the purpose of grading (3). The paper raises further questions, tensions, and issues around: 1) the purposes of peer assessment, 2) contextual factors that impact on raters and ratings, and 3) standards of professional behaviour, which we will further explore.

The focus in higher education environments is increasingly turning to the use of peer assessment for learning, that is, in a formative manner to improve student learning of understanding or skill (4,5). It has been hypothesised that it is not so much the receiving of feedback, but rather the learning process of making an evaluative judgement and constructing feedback, that is the crucial, learning component (6). Even if we accept [authors’] position that peer assessments should be used in a formative, developmental fashion, we must still take into consideration the contextual factors under which they are implemented.

Such contextual factors include the purpose of the assessment and how it will influence students’ behaviour. The primary aims of the authors’ research were 1) “to cue desired behaviour changes as the student considered their own performance profile” and 2) “to cue desirable rating behaviour changes as the student considered experientially, from a recipient’s perspective, the meaning and significance of the rating response that they are providing for others” (3). [authors]’ intent for the instrument highlights the reciprocity inherent in such a peer assessment, especially when there are multiple opportunities for peer assessment. Separating the need to assess from being assessed is nigh impossible, and therefore being assessed will influence the assessment. In this case, the relatively safe nature of the assessments (within a familiar group, and not for summative purposes) are less likely to produce inaccurate scores (e.g. due to collusion, peer pressure, or rating of strangers). Van Gennip, Segers and Tillema (7) highlight the importance of optimising interpersonal variables (i.e. psychological safety, trust, value diversity, and interdependence) in peer assessment, because peer assessment is a social action that is part of a peer collaborative learning process.

In addition to raising awareness of such influences on peer assessment, authors’ paper also shines light on questions pertaining to standards of professional behaviour – “what is professional behaviour?”, and “how does one measure professional behaviour?” or indeed, “can a set of criteria be constructed to form a finite set of professional competence?” [authors] adapt a questionnaire from Papinczak (8), appropriate for PBL and other tutorial settings. What “professional” looks like on a clinical placement might have additional dimensions and distinctive standards of behaviour. If one purpose of a medical program is engendering students’ professional behaviour, it is perhaps less surprising that scores become less varied in the second year. We might reasonably expect that
students will become more merged and transform towards “professional” over time, with fewer students scoring poorly (9).

By raising these questions, authors’ study presents implications for the validity of assessments of professional behaviour more generally. As they indicate, “what has not been fully explored, is whether inferences about the behaviour of a student within a PBL group can be generalised to make claims about their behaviour relative to the whole cohort” (3). Should we be considering professional behaviours as relative qualities comparable to others in the same cohort, or as absolutes that can be considered acceptable and unacceptable? We hazard to suggest that there is a definite set of standards for professional behaviour in the moment-to-moment interactions that were observed although they may also change over time (e.g., as it becomes more common to use social media to communicate with patients). This is, essentially, the debate for normative or criterion based standards rehashed in the context of professional behaviour.

Given the multiple purposes of peer assessment, careful consideration of the design and set-up is required and should be driven by the desired outcomes. The assessment of professional behaviour might best be achieved through peers: in professional practice, it is essentially our peers that we are beholden to, through professional bodies and associations. Since peer assessments are inherently a social undertaking, investigating social perspectives is paramount to advance our understanding of how and why we should carry out peer assessments in the future. [authors] certainly raise relevant questions around this line of argument - reminding us to look deeper into the opportunities and challenges that peer assessment brings to the learning of professional behaviour in medical education.

References:

3. [authors’ paper for which this commentary is written]