Article type : Commentary

Corresponding author mail id: scarffc@student.unimelb.edu.au

Towards a greater understanding of narrative data on trainee performance.

In a research paper in this issue of *Medical Education*, Wilby et al (1) present a study on the interpretation of narrative data generated by assessors in a final year OSCE for pharmacy students in Qatar. The authors collated the written comments on the performance of a sample of nine students (of good, average and poor performance standard) across all 6 of the OSCE stations and enlisted expert-assessors to review them, using think-aloud procedures and verbal protocol analysis. The overall purpose was to explore how expert-assessors understand the narratives written by others in the setting of a standardized performance assessment. In other words, how we seek to reconstruct a learner's unwitnessed performance based on descriptions provided by others.

The authors note that narrative assessment data are increasingly being called upon for use across many assessment formats. They and others note the value of this type of information: both its ability to contextualise numerical data (2) and to go further, providing useful information beyond the confines of the ratings to offer a more global view of the trainee's performance (3). While some trainees may prefer comments to ratings (4), both have their strengths and weaknesses. For example, some studies have shown that qualitative data can reveal increased numbers of poor performers compared with quantitative data (5) and that it is less susceptible to issues such as grade inflation (6). However, the work required to provide such narrative data and to communicate it to the learner is not insignificant and cognitive, motivational and normative influences are relevant.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/MEDU.13940](https://doi.org/10.1111/MEDU.13940)

The human resources appraisal literature reminds us of the profound difference between “ticking a box” to reflect the level of an individual’s performance and actually composing text about it (7) and the significant cognitive workload in the OSCE setting in particular has been recognised (8). Though describing other assessment formats, it is recognised that at times, assessors are not motivated to provide comments and avoid doing so (2,9), even when they are a required part of an assessment protocol (10). Better understanding of the reasons for this is the focus much research – including my own.

Despite the value of narrative data, issues can arise if the two components of an assessment – the numerical and written narrative parts – do not tell a similar or consistent story about a trainee’s level of performance (11). The accurate interpretation of narrative comments is key if they are to achieve their authors’ aim of providing additional, useful information to trainees. Ginsburg et al. have shown that both trainees and assessors are able to understand the level of performance of a trainee based on comments made in the narrative section of assessment forms (4,12), but it is unknown if all can ‘crack the code’ to the same extent. The time and effort it takes to do so is also relevant (2). So, while composing narrative data has its own level of difficulty, interpreting it also has challenges. Wilby et al. seek to address this, by increasing our understanding of how assessors make meaning of such narrative comments.

They found that expert-assessors used two strategies (comparing and contrasting and the forming of mental images) and three perspectives (the student’s, the examiner’s and that of the professional) when interpreting and understanding the narrative comments in the study. This finding provides an important perspective on the potential origins of variability in assessment – the assessors’ own interpretations of student performance – and the richness of the assessment message that can, therefore, ultimately arise. The authors argue the way forward is to appreciate the variation and nuance arising from different interpretations rather than aiming to eliminate it with more rigorous assessor training or simplifying data further. They note the potential benefit of having an awareness of methods assessors use when judging a trainee, and how this could be applied to their training, such as in the avoidance of normative judgements. My colleagues and I echo their (and others’) call for the use of multiple assessors over time and in committee format, especially in relation to progress decisions. We have observed assessors call for the support of their colleagues with such decisions, especially when the performance outcome is a negative one with likely emotional load for both recipient and messenger. While Wilby et al’s future work aims to build on their

findings and compare interpretations to outcome decisions, it is also important to consider the trainees' perspectives on the information received. In a recent review conducted with colleagues (13), the trainee perspective of the assessment messages delivered in the clinical performance assessment setting was examined. The finding that some trainees do not value these messages, and that the perceived value is heavily dependent on factors in their environment, their assessor and themselves, may well have some lessons for the OSCE context also.

It is important that we listen to and, where possible, address the needs of the learners for whom the assessments are designed. Determining the trainees' strategies for making meaning of written narrative comments would be an important additional area to study to shed light on how they interpret the data and understand their level of performance and future needs. How this relates to the numerical information they receive and their handling of seemingly contradictory messages would be valuable to explore to avoid resultant confusion and mistrust in assessment and potential disregarding of the assessment message. After all, the successful conveyance of this message is at its heart an interpersonal transaction, and these are eternally prone to misunderstanding.

Acknowledgements – Steve Trumble, Margaret Bearman, Neville Chiavaroli.

Pull-out quotes

1. narrative assessment data are increasingly being called upon for use across many assessment formats
2. the work required to provide such narrative data and to communicate it to the learner is not insignificant
3. The accurate interpretation of narrative comments is key if they are to achieve their authors' aim of providing additional, useful information to trainees
4. The authors argue the way forward is to appreciate the variation and nuance arising from different interpretations rather than aiming to eliminate it with more rigorous assessor training or simplifying data further
5. It is important that we listen to and, where possible, address the needs of the learners for whom the assessments are designed

1. Wilby KJ, Dolmans DH, Austin Z, Govaerts MJ. Assessors' interpretation of narrative data for communication skills from a summative OSCE. *Med Educ*. 2019;
2. Richards SH, Campbell JL, Walshaw E, Dickens A, Greco M. A multi-method analysis of free-text comments from the UK general medical council colleague questionnaires. *Med Educ*. 2009;43(8):757–66.
3. Ginsburg S, Gold W, Cavalcanti RB, Kurabi B, McDonald-Blumer H. Competencies “Plus”: The Nature of Written Comments on Internal Medicine Residents' Evaluation Forms. *Acad Med*. 2011;86(10):S30–4.
4. Ginsburg S, van der Vleuten CPM, Eva KW, Lingard L. Cracking the code: Residents' interpretations of written assessment comments. *Med Educ*. 2017;51(4):401–10.
5. Durning SJ, Hanson J, Gilliland W, McManigle JM, Waechter D, Pangaro LN. Using Qualitative Data From a Program Director's Evaluation Form as an Outcome Measurement for Medical School. *Mil Med*. 2010;175(6):448–52.
6. Battistone MJ, Pendleton B, Milne C, Battistone ML, Sande MA, Hemmer PA, et al. Global descriptive evaluations are more responsive than global numeric ratings in detecting students' progress during the inpatient portion of an internal medicine clerkship. *Acad Med*. 2001;76(10 Suppl):S105–7.
7. Brutus S. Words versus numbers: A theoretical exploration of giving and receiving narrative comments in performance appraisal. *Hum Resour Manag Rev*. 2010;20(2):144–57.
8. Byrne A, Tweed N, Halligan C. A pilot study of the mental workload of objective structured clinical examination examiners. *Med Educ*. 2014;48(3):262–7.
9. Plymale MA, Donnelly MB, Lawton J, Pulito AR, Mentzer RM. Faculty evaluation of surgery clerkship students: important components of written comments. *Acad Med*. 2002;77(10 Suppl):S45–7.
10. Sebok-Syer SS, Klinger DA, Sherbino J, Chan TM. Mixed Messages or Miscommunication? Investigating the Relationship Between Assessors' Workplace-Based Assessment Scores and Written Comments. *Acad Med*. 2017;92(12):1774–9.
11. Mulliner E, Tucker M. Feedback on feedback practice: perceptions of students and academics. *Assess Eval High Educ*. 2017;42(2):266–88.

12. Ginsburg S, Regehr G, Lingard L, Eva KW. Reading between the lines: Faculty interpretations of narrative evaluation comments. *Med Educ*. 2015;49(3):296–306.
13. Scarff CE, Bearman M, Chiavaroli N, Trumble S. Trainees' perspectives of assessment messages: a narrative systematic review. *Med Educ*. 2019 Mar;53(3):221–33.

Author Manuscript

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Scarff, CE

Title:

Towards a greater understanding of narrative data on trainee performance

Date:

2019-08-11

Citation:

Scarff, C. E. (2019). Towards a greater understanding of narrative data on trainee performance. MEDICAL EDUCATION, 53 (10), pp.962-964. <https://doi.org/10.1111/medu.13940>.

Persistent Link:

<http://hdl.handle.net/11343/286262>