

Team Assessment Behaviour (TAB) as an assessment tool: A critical evaluation

Milind Pant, Prabhu N Nesargikar and Daniel M Cocker

Abstract

The concept of evaluation of professionalism has gained momentum in the UK in recent years, following concerns about poorly performing doctors. Multi source feedback was designed with an aim to evaluate the professional behaviour of a doctor, which is generally hard to assess as it is multivariate. In the UK, Team Assessment Behaviour (TAB) is one of the assessment tools used under the umbrella term of Multisource Feedback (MSF) in postgraduate and undergraduate medicine. As TAB is a relatively new tool which has been developed and tested within the UK, we have described the process of TAB along with defining the barriers that may prevent it and the MSF from being an effective tool.

Introduction

Good Medical Practice describes the professional behaviour expected of doctors and advocates that it should be taught as well as assessed. GMC's Good Medical Practice specifies the standards of team working, communication skills, accessibility and trustworthiness in relation to professional behaviour¹.

Ramsden² stated that: "the students will learn what they think they will be assessed on, not what is in the curriculum, or even on what has been 'covered in class'." Hence, if the intended learning objectives are to improve professional behaviour and team working in a trainee doctor, then a tool that assesses these characteristics along with providing suggestions for improvement is critical.

Multisource Feedback (MSF) is a formative assessment tool that was designed to assess professional behaviour and attitudes, with the aim of continually improving an individual's team working. Team Assessment of Behaviour (TAB) is an assessment tool for MSF (Appendix 1), and is one of the 2 assessment tools used to assess professional behaviour of foundation doctors in training in the UK, the other being the Mini-Peer Assessment Tool (mini-PAT)³.

Our aims for this article were to evaluate whether the format of TAB allowed the MSF process to occur as originally intended, along with looking at possible barriers that may have to be overcome to make it an effective tool. A literature review was carried to appraise the present evidence regarding TAB as an assessment tool and only studies which had relevance to MSF/TAB were included in the study. As TAB is a relatively new tool, there were very few papers' exploring this in depth which was a major limitation.

Hence in the barriers section, we have discussed the possible obstacles to the whole MSF process rather than just TAB.

Background

The concept of MSF was originally developed by industrial organisations and has been used in postgraduate medicine in USA for assessment of professional behaviour since the 1990's^{4,5}. Ramsey⁶ suggested that it is feasible to obtain assessments from professional associates of practicing physicians in areas such as clinical skills, humanistic qualities, and communication skills.

TAB was developed by educationalists and senior doctors at the West Midlands deanery^{7,8}, and has undergone extensive field testing among 171 trainees, with analysis of received feedback from 1378 assessments across four different hospitals in the West Midlands. It is currently being used in the West Midlands Deanery for the multisource feedback of foundation programme trainees.

Process of TAB

The primary aim of TAB as an assessment tool is to identify trainees whose professional behaviour does not meet GMC requirements for good medical practice, so that appropriate action may be taken and also to compliment those trainees who receive good reports. For TAB, up to 10-15 multi-disciplinary colleagues of a doctor assess his/her workplace behaviour. It assesses four domains of professional behaviour: professional relationship with patients, verbal communications, team-working and accessibility.

It is the initiative of the trainee to distribute at least 15 TAB forms to peers of their own choice and a minimum of 10 completed forms are required to be returned. The raters should

include at least three other doctors including a consultant supervisor and at least five allied healthcare professionals. It's the responsibility of the educational supervisor to collate and summarise these forms, identifying perceived weaknesses, offering them feedback and directed learning objectives to address any issues.

Critique of TAB:

1. Validity and reliability of TAB:

Validity is a demonstration that a particular instrument can in fact measure what it purports to measure⁹. TAB portrays face as well as content validity as it assesses areas identified by the GMC¹ for good professional behaviour. It is shown to be capable of identifying problem behaviour in trainees, which was the primary aim of the tool⁷. This tool appears to have good construct validity as it is testing trainee's behaviour in real life situations.

It is difficult to define the predictive validity of any tool, more so in a formative assessment tool. The developers of TAB have not tested for concurrent validity. The concurrent validity for TAB could have been tested by using another MSF tool in addition to TAB for some participants during the field-testing for reliability and validity of the tool.

A reliable instrument for a piece of research will yield similar data from similar respondents over a period of time⁹. Whitehouse⁷ demonstrated in the pilot study that TAB had intra-observer reliability and inter-observer variability. For inter-observer variability, the Royal College of Psychiatrists compiled a map of assessment programmes against good medical practice domains and considered it appropriate for assessing four domains: good clinical care, working with colleagues, probity, and health¹¹.

2. Raters

Ramsey et al⁵ concluded that, with MSF, 10-11 responses per physician were necessary to achieve a generalisability coefficient of 0.70. Wood et al¹¹ concluded that eight raters were sufficient for a representative score in their study on Obstetrics & Gynaecology trainees in the UK. Obviously more raters would lead to better coefficient and more generalisable results. TAB presently advocates at least ten raters to achieve reliable results.

3. Feasibility

TAB has four domains and a three point rating scale which are relatively easy to understand and complete. There is no training required for raters and usually takes less than five minutes per assessment per assessor. The paper-based system demanded considerable administrative resources, and therefore a web-based TAB assessment form was successfully piloted in the West Midlands⁸.

4. Trainer & trainee's evaluation of the process

The evaluation of TAB as an assessment tool was done by Whitehouse et al⁸ as a part of their field assessment of TAB. The assessors and trainers found the process practical, valuable and fair. 76% of the trainees who responded to the questionnaire felt that it was a useful addition to the assessment of the SHO's. The educational supervisors had mixed views, with 77% of them finding out nothing new about the trainees.

5. Scoring system

Instead of a Likert scale, the TAB employs broad boxes which offer the rater a choice of giving specific feedback under each domain. This is more helpful to a summariser/assessor than mini-PAT, where a scoring scale is not substantiated by relevant feedback.

This assessment tool does not assess clinical performance and one could argue that there could have been more than four domains in order to include other areas of performance, such as clinical skills. However having just four domains can reduce the impact of the halo effect. The halo effect¹² can be defined as a rater making an overall judgement of the trainee and scoring the whole form accordingly rather than considering each domain separately. This could be a potential advantage of TAB over mini-PAT where there is no opportunity to provide specific feedback in individual domains, but rather utilises an overall action plan, which may lead to concentrating single element deficiencies while masking other shortcomings.

6. Patient Outcomes

The MSF process in itself doesn't bear direct consequences on patients' management, but can help the doctor improve his professionalism, which is critical element of good medical practice¹. An improvement in a doctor's behaviour secondary to the appraisal-feedback process can indirectly improve and contribute to patient management and satisfaction.

Potential barriers:

1. Choice of rater:

The trainee has the choice of raters, except for the supervising consultant, who must be involved in the process. This is a potential area of difficulty, as the trainee might pick individuals who are more sympathetic to their cause, or who cannot comment much on their interpersonal behaviour, thus leading to skewed results. Kuzmits¹³ showed that both rater and those being rated needed to be trained to make the rating and feedback process more effective, but this is not felt necessary for TAB. There is also a potential conflict between being a trainer and assessor, and this conflict might lead to clouding of judgement¹⁴, which can be addressed by having different assessors and trainers, but considering the service pressures in the NHS, this may not always be possible.

2. Feedback:

The value of the MSF process can be limited by the quality of feedback provided to the trainee at the end of the process and depends on the relationship of the trainee and the supervisor¹⁵. Evidence shows that non-specific feedback does little to change performance¹⁶. Whitehouse⁸ concluded in their initial study that TAB was able to produce descriptive feedback that was more specific and helpful than existing MSF tools. Assessors completing TAB may not give specific comments or feedback although they are instructed to give some details especially if they chose the rating of "some concern" or "major concern". Bret and Atwater¹⁷ have shown that negative feedback can discourage individuals, and they can even react in anger.

Feedback about performance must be descriptive and specific if it is to be helpful to trainees¹⁸. Hence it is the responsibility of the supervisor to give feedback which is relevant and helpful, along with creating an action plan with the trainee to address any perceived deficiencies. It must also be stressed that good performers need to be complimented, and encouraged to continue to do so.

3. Training the raters and the supervisor:

The raters also need to be educated in the process otherwise they may not give reliable views about the trainee^{13,19}. From the rater's point of view, it would be beneficial if they were to give specific comment, in order that more relevant feedback can be provided.

Most importantly, the educational supervisors who provide the feedback to the trainee can make a difference in constructing agreed learning objectives, and not demoralise the trainee with negative feedback. Kaplan²⁰ also noted that negative feedback can demotivate individuals. Giving face to face negative feedback can be a daunting task and supervisors, if not trained in giving negative feedback may dislike doing it, and so not give properly constructed feedback. The provision of training, however, has time as well as cost implications.

Conclusions

TAB continues to be part of the national foundation programme curriculum and, used correctly, can serve its purposes both as a screening tool and also for the trainees to use the feedback provided to improve their interpersonal behaviour when needed. Evaluation and quality assurance of this assessment tool should be an ongoing process. More field work in relation to assessment of behaviour in relation to TAB is needed. Patient feedback could be included in TAB which can make it a more reliable tool for assessing a doctor's behaviour.

Further qualitative studies to explore the views and experiences of trainees can help to understand the barriers and attempt to improve the usefulness of the process for the trainees. There are

several potential barriers which can subvert the process of MSF by using TAB, and these need to be addressed to make the assessment process more effective and efficient.

COMPETING INTERESTS

None Declared

AUTHOR DETAILS

MILIND PANT, MRCPsych, Psychiatric Teaching Fellow, Keele School of Medicine, UK
 PRABHU N NESARGIKAR, MRCS, Surgical Teaching Fellow, Keele School of Medicine, UK
 DANIEL M COCKER, MRCS, Surgical Teaching Fellow, Keele School of Medicine, UK
 CORRESPONDENCE: MILIND PANT, 26, Prince town Close, Meir Park, Stoke on Trent, ST3 7WN, UK
 Email: milindpant@hotmail.com

REFERENCES

1. Good Medical Practice: General medical council London (2001).
2. Ramsden P. Learning to teach in higher education. 1992. London Routledge Press.
3. MMC, National Foundation programme, UK
4. Woolliscroft J, Howell J, Patel BP, Swanson DB. Resident-patient interactions: The humanistic qualities of internal medicine residents assessed by patients, attending physicians, program supervisors, and nurses. *Academic Medicine* 1994; 69: 216-224.
5. Ramsey PG, Carline JD, Blank LL, Wenrich MD. Feasibility of hospital-based use of peer ratings to evaluate the performances of practicing physicians. *Academic Medicine* 1996; 71: 364-70.
6. Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA* 1993; 269: 1655-60.
7. Whitehouse A, Hassell A, Wood L, Wall D, Walzman M, Campbell I. Development and reliability testing of TAB a form for 360 degree assessment of Senior House Officers' professional behaviour, as specified by the GMC. *Medical Teacher* 2005; 27:252-8.
8. Whitehouse A, Hassell A, Bullock A, Wood L, Wall D. 360 degree assessment (multisource feedback) of UK trainee doctors: Field testing of team assessment of behaviours (TAB). *Medical Teacher* 2007; 29:171-176.
9. Cohen L, Manion L (2000). *Research Methods in Education: 2000:5th Edition*: Routledge Press.
10. Royal College of Psychiatrists London UK.
11. Wood L, Wall D, Bullock A, Hassell A, Whitehouse A, Campbell I. 'Team Observation': A six year study of the development and use of multi-source feedback (360 degree assessment) in obstetrics and gynaecology training the UK. *Medical Teacher* 2006; 28: 177-84.
12. Thorndike, E.L. A constant error in psychological ratings. *Journal of Applied Psychology* 1920; 4:25-29.
13. Kuzmits FE, Adams AJ, Sussman L, Raho LE. 360-feedback in health care management: a field study. *The Health Care Manager* 2004; 23: 321-8.
14. Bray L, Nettleton P. Assessor or mentor? Role confusion in professional education: *Nurse Education today* 2007; 27: 848-55.
15. 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: 199-218.
16. Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback: perceptions of credibility and usefulness. *Medical Education* 2007; 39: 497-504.
17. Bret and Atwater. *The Journal of Applied Psychology* 2001; 86: 930-42.
18. Mohanna K, Wall D, Chambers R. *Teaching Methods Made Easy: A manual for Health Professionals*, 2004; 2nd edition; Abingdon, Radcliffe Medical Press.
19. Abdulla A. A critical analysis of mini peer assessment tool (mini-PAT): *J R Soc Med* 2008; 101: 22-26
20. Kaplan R. 360 degree feedback PLUS: boosting the power of co-worker ratings for executives. *Human Resource Management* 1993; 32: 299-314