

# Team assessment of behaviour: a high stakes assessment with potential for poor implementation and impaired validity

**Authors:** Andrew Whitehouse,<sup>A</sup> Laura Higginbotham,<sup>B</sup> Kamal Nathavitharana,<sup>C</sup> Baldev Singh<sup>D</sup> and Andrew Hassell<sup>E</sup>

## ABSTRACT

**Team assessment of behaviour (TAB) is the multi-source feedback assessment of professional behaviours that all UK foundation doctors must engage in twice during their two-year programme. TAB can identify the few underperforming trainees and provide feedback to consolidate the good practice of most. For optimum validity, TAB must be undertaken by a range of assessors, as specified in the national UK Foundation Programme curriculum. This study reports an audit of invalid TAB submissions over a three-year cycle in the West Midlands' Foundation Programme. In 2010, large numbers of TABs were invalid, owing to an incorrect selection or number of assessors. Introduction of validity checking before sign-off greatly improved the numbers of valid assessments in 2011. This was partially sustained in 2012. Assurance of assessment validity is important to ensure delivery of appropriate constructive feedback and to allow early detection and remediation of signs of poor professional behaviours in foundation doctors.**

**KEYWORDS:** TAB, assessment, validity, multi-source feedback, foundation programme

## Introduction

Doctors in training are increasingly subject to assessment in their workplace. The UK Foundation Programme curriculum requires all trainees to undertake at least one annual multi-source feedback assessment using the 360° assessment tool, team assessment of behaviour (TAB). The tool is a validated assessment designed to identify trainees with substandard professional behaviour in the areas of team working, reliability/

accessibility, verbal communication skills and maintaining trust, and also to provide them with specific written feedback, which usually serves to recognise and consolidate the professionalism exhibited by the great majority. All foundation doctors must engage in this assessment, at least once in each year, to graduate from foundation and move on to specialty training. It is therefore a high stakes assessment, which must be carried out to curriculum standards. It is important that it is undertaken with robust validity. This paper examines the validity of TAB implementation over three years within a single deanery.

## Valid assessment

Based on published research<sup>1–3</sup> there are several elements of TAB assessment that must be delivered correctly to optimise validity, and that were listed in the 2010 curriculum.<sup>4</sup> 'For each assessment, the foundation doctor should nominate 15 raters. A minimum of 10 returns is required. No other foundation doctor can be a rater.'

The recommended mix of raters/assessors in the curriculum is as follows:

- > 2–8 doctors more senior than foundation (F)2, including at least one consultant or GP principal
- > 2–6 senior nurses (band 5 or above)
- > 2–4 allied health professionals
- > 2–4 other team members including ward clerks, secretaries and auxiliary staff.

The 2012 version of the curriculum<sup>5</sup> modified the required range of assessors to:

*... at least two of each of the following:*

*Doctors more senior than F2 including at least one consultant or GP principal*

*Senior nurses (band 5 or higher)*

*Allied health professionals*

*Other team members, e.g. ward clerks, secretaries, auxiliary staff.*

Further guidance is provided on the action to be taken if concerns about a trainee are raised in TAB. After one-to-one feedback with an educational supervisor, the TAB summary reports should be released to the trainee. The summary should include all the anonymous ratings and verbatim comments.

This 'release' triggers the closure of the TAB assessment, which allows the e-portfolio to commence a follow-up TAB

**Authors:** <sup>A</sup>associate postgraduate dean, Health Education West Midlands, Birmingham, UK; <sup>B</sup>revalidation support officer, Health Education West Midlands, Birmingham, UK; <sup>C</sup>consultant paediatrician, Health Education West Midlands, Birmingham, UK; <sup>D</sup>consultant physician in diabetes and endocrinology, Health Education West Midlands, Birmingham, UK; <sup>E</sup>consultant rheumatologist, Haywood Hospital, Stoke-on-Trent, UK, and director of undergraduate programmes, School of Medicine, Keele University, Staffordshire, UK

round if this is needed either in the same year, or, for all F1s, when it is needed in the following, F2 year.

## The 2010–2011–2012 audit of TAB in the West Midlands Deanery

### Method

In the deanery's five foundation schools all trainees undertake their first, and usually only, annual TAB assessment between mid-October and late November. The outcomes are registered on the e-portfolio of each trainee, which can be reviewed by administrative leads and clinical and educational supervisors.

This audit was undertaken to check that certain elements of the assessment for our trainees had been properly observed, allowing reassurance that the assessments had been valid.

The elements of the assessment selected as the standards to be checked were as follows:

- a minimum of 10 assessments must be recorded
- no other foundation doctor may be a rater
- assessors must include at least one consultant or GP principal.

TAB data from the three years of TAB assessments were reviewed through the deanery administrators' access to the e-portfolio database for all our foundation trainees.

After the 2010 analysis, new advice was issued to all directors of medical education (DMEs), foundation directors and post-graduate centre (PGC) staff on how the validity of each TAB should be checked before the summary was released to the trainee. It was specified that the validity of every TAB should be checked by a suitably trained staff member before the summary's release.

Summaries of valid TABs could be released by PGC staff, provided that no concern had been recorded by any assessor. If the assessment was invalid, or if any concern had been recorded, then one-to-one feedback by a DME or educational supervisor should take place and a repeat TAB be set up for the next placement, as the curriculum requires.

After analysis of the 2012 TAB round, a questionnaire was sent to all PGCs in our deanery to establish whether the process for validity checking before summary release, which was prescribed in 2011, had continued effectively in 2012.

### Results: 2010

Table 1 shows the total numbers of TABs performed and the numbers and percentages of invalid TABs for each of the three years studied.

As the part of Table 1 for the initial 2010 analysis shows, 630 F1s and 590 F2s undertook a TAB assessment during the reference period. Of these only 437 (69.4%) and 380 (64.4%), respectively, received valid assessments by all the three standards considered. Table 1 shows that many trainees received fewer than 10 assessments and, of those who did achieve the minimum 10, many had selected foundation trainee assessors. Others had selected foundation trainee assessors, but, as they had at least 10 acceptable non-foundation assessors anyway, their assessments were valid. Some trainees had, however, chosen several other foundation trainees to assess them, in some cases as many as five or six. As for the standard that at least one assessor must be a consultant or GP principal, we found that almost a quarter of trainee TAB assessments did not include such an assessor.

### Results: 2011

After the change of validity checking and release policy, the 2011 audit confirmed a gratifying improvement in the rates of valid assessments, as shown in the part of Table 1 dealing with 2011. This time >95% of trainees underwent valid TAB assessments.

### Results: 2012

The 2012 audit revealed a significant relapse in the numbers of valid assessments, albeit not to the poor levels of 2010 (see the part of Table 1 dealing with 2012). The questionnaire

**Table 1. Completed TABs among West Midlands Deanery foundation doctors by year and grade, showing the total number that were invalid (number (percentage of total completed)) and also the categorisation of invalidity (number (percentage of total completed) (percentage of those invalid)).**

Year	Grade	Completed TAB	Invalid TAB numbers (%)			
			Total	<10 assessors	<10 non-foundation assessors	No senior doctor
2010	All	1,220	403 (33)	302 (25) (75)	101 (8) (25)	286 (23) (71)
	FY1	630	193 (31)	139 (22) (72)	54 (9) (28)	144 (23) (75)
	FY2	590	210 (36)	163 (28) (78)	47 (8) (22)	142 (24) (68)
2011	All	1,286	70 (5)	24 (2) (34)	22 (2) (31)	39 (3) (56)
	FY1	643	34 (5)	12 (2) (35)	10 (2) (29)	23 (4) (67)
	FY2	643	36 (6)	12 (2) (33)	12 (2) (33)	16 (3) (44)
2012	All	1,257	326 (26)	153 (12) (47)	22 (2) (7)	233 (19) (71)
	FY1	633	170 (27)	59 (9) (35)	12 (2) (7)	132 (21) (78)
	FY2	624	156 (25)	94 (15) (60)	10 (2) (6)	101 (16) (65)

An invalid TAB could have more than one category of invalidity. FY = foundation year; TAB = team assessments of behaviour.

about TAB release methods used in 2012 was sent to 17 local education providers (LEPs), of whom 12 responded. Of these 12, although reporting that they still tried to check all assessments before release, some had indeed been more diligent in this during the 2011 round than in the 2012 one. This was also the first year in which all our foundation trainees had a single educational supervisor for the whole year, so some supervisors were new and less familiar with the rules on release than their predecessors.

## Discussion

TAB was devised, validated and first used in the West Midlands deanery in 2004. Its design, in particular its simple 3-point rating scale,<sup>6</sup> improves its sensitivity to detect trainees who exhibit professional behaviours that cause concern to colleagues. It also prompts written feedback from assessors, describing the detail of both concerning and admired behaviours, and thus TAB produces much supportive feedback to most young doctors who exhibit proper professional behaviours. However, to optimise the robustness of TAB in identifying trainees with substandard behaviour and to provide feedback to all trainees, the assessment must be carried out in accordance with the evidence-based guidance. Earlier work has shown that, with the correct profile of assessors, and a minimum of 10 assessments, TAB is more likely to identify trainees whose professional behaviour has caused concern. The specification in the foundation curriculum that there must be at least two senior nurses and two senior doctors among the assessors is based on research, which concludes that these are the likeliest professional groups to identify concerns. The spread of other professional groups required allows the widest range of potentially positive feedback to be harvested for each trainee.

It may seem surprising that, even in a deanery with long experience of using TAB, a poor level of compliance with prescribed rules has been found. It is possible that the low validity rate for 2010 was due to changes in the professional group profile required for the assessment, which the foundation curriculum had included in its revised 2010 edition. Before 2010 there was no embargo on fellow foundation doctors assessing their peers. This change was made after a study that revealed a very low number of concern ratings in TABs from fellow foundation doctors, reducing the tool's sensitivity to identify poor professional behaviours.<sup>3</sup>

We have no reason to believe that TAB use in other deaneries complies more or less well with the curriculum validity rules, and we recommend that all foundation schools, for their own quality management purposes, undertake a similar analysis, using e-portfolio review, to make sure that the curriculum specifications on this assessment are being met.

We also firmly advise that TAB summary release to a foundation trainee should not take place in any LEP until the validity of the assessment, with respect to the number of returns and the professional group profile of assessors, has been confirmed by a reliable person. This can be undertaken through each foundation training PGC, as our 2011 results show, where the administrative staff, trained and led by foundation school directors and DMEs, can review each summary and check that each contains the required number and professional group of

assessors. Although it has normally been advised that all TAB summaries are fed back to the trainee personally, our previous experience has been that not all educational supervisors actually do this, sometimes meaning that valuable positive feedback failed to reach the trainee. We therefore now advise that PGC staff may release any valid assessment summaries that contain only positive ratings and comments directly to trainees. Summaries that are invalid or contain concern ratings or comments must still be fed back in one-to-one discussion with an educational supervisor, DME or programme director.

The required assessor groupings for TAB are being considered in the 2016 UK Foundation Programme curriculum revision. This may lead to changes in the e-portfolio, which will simplify input of an assessor's professional role. Ideally the TAB programme on the foundation e-portfolio would exclude ineligible assessors from submitting feedback and prevent 'summary release' without a fully validated assessment.

Improving the TAB process will facilitate routine foundation school analysis of this widely valued assessment. In addition it may potentially free-up time for senior clinicians to spend on delivering feedback, rather than taking part in the validity checking process or trying to keep abreast of detailed curricular changes.

Multi-source feedback continues to offer considerable potential in identifying trainees with suboptimal professional behaviours and in providing helpful feedback to all trainees. It is, however, essential that it be implemented according to best evidence. We would suggest that this validity analysis be repeated in other deaneries and recommend it as a simple measure of best practice. At the same time, we believe that there is scope for further work to enhance the usefulness of multi-source feedback in postgraduate medical training. ■

## References

- 1 Whitehouse A, Hassell A, Wood L *et al.* Development and reliability testing of TAB, a form for 360° assessment of senior house officers' professional behaviour, as specified by the General Medical Council. *Med Teacher* 2005;27:252–8.
- 2 Whitehouse A, Hassell A, Bullock A, Wood L, Wall D. 360 degree assessment (multisource feedback) of UK trainee doctors: field testing of 'TAB'. *Med Teacher* 2007;29:171–6.
- 3 Bullock A, Hassell A, Markham W, Wall D, Whitehouse A. How ratings vary by staff group in a multi-source feedback assessment of junior doctors. *Med Educ* 2009;43:516–20.
- 4 UK Foundation Programme Office. *The UK foundation programme curriculum, 2010*. Available online at [www.foundationprogramme.nhs.uk/pages/home/training-and-assessment](http://www.foundationprogramme.nhs.uk/pages/home/training-and-assessment) [Accessed 30 October 2011].
- 5 UK Foundation Programme Office. *The UK foundation programme curriculum, 2012*. Available online at [www.foundationprogramme.nhs.uk/index.asp?page=home/keydocs#c&rg](http://www.foundationprogramme.nhs.uk/index.asp?page=home/keydocs#c&rg) [Accessed 6 August 2014].
- 6 Hassell A, Bullock A, Whitehouse A *et al.* Effect of ratings scales on scores given to junior doctors in multi-source feedback. *Postgrad Med J* 2012;88:10–14.

**Address for correspondence:** Dr A Whitehouse, Health Education West Midlands, St Chads House, Hagley Rd, Birmingham, UK.  
**Email:** [a.b.whitehouse@bham.ac.uk](mailto:a.b.whitehouse@bham.ac.uk)