Situational judgment tests reliably measure professional attributes important for clinical practice

Katie Petty-Saphon¹
Kim A Walker²
Fiona Patterson³,⁴
Vicki Ashworth⁴
Helena Edwards⁴

¹Medical Schools Council, London, UK; ²United Kingdom Foundation Programme Office, Birmingham, UK; ³Department of Psychology, University of Cambridge, Cambridge, UK; ⁴Work Psychology Group, Derby, UK

Abstract: Over the course of more than 40 years, international research has consistently shown situational judgment tests (SJTs) to be a reliable and valid selection method for assessing a range of professional attributes. However, SJTs still represent a relatively new selection method within the medical profession, and as such it is to be expected that applicant reactions will vary. In this Expert Opinion piece, we respond to Najim et al’s article “The situational judgement test: a student’s worst nightmare” by highlighting three key clarifications. We outline that 1) the UK Foundation Programme’s SJT deliberately measures only a subset (five) of the nine professional attributes important for the role of Foundation Trainee doctor, 2) these attributes are measured in addition to academic attainment, and 3) the SJT represents a cost-effective approach to selection rather than attempting to interview approximately 8,000 candidates each year, which would be logistically impossible. We present these points to inform future research and encourage debate, and conclude that the SJT is an appropriate and fair measurement method to be used as one part of selection to the UK Foundation Programme.

Keywords: situational judgment test, reliability, validity, fairness, candidate perceptions

Article

International research consistently shows situational judgment tests (SJTs) to be a reliable and valid selection method for a range of professional attributes (such as integrity, empathy, and teamwork) important in any health care profession.¹ Given this evidence, SJTs have received a great deal of attention both in the academic literature and through commentary from stakeholders including trainees, which is unsurprising given their increasing use within medical selection. While there is over 40 years of research evidence on SJTs, it is acknowledged that, as a relatively new method within medical selection, applicant reactions are not always positive. We thank Najim et al² for raising some important concerns in relation to perceptions of the UK Foundation SJT. In a recent review of the research evidence,³ SJTs are shown to be a reliable and valid selection method compared to many other tools, including personal statements (or white space questions) on application forms. In response to Najim et al’s letter, we highlight three key clarifications to inform further debate.

1. Najim et al criticize the test specification for the Foundation SJT and coverage of the attributes measured: “the authors cannot comprehend how the written SJT fairly and accurately tests these attributes.” This view assumes that the SJT attempts to measure all nine of the professional attributes required for the role of foundation doctor identified through a job analysis study.³ The job analysis report and all permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).
subsequent communications regarding the SJT (including a monograph and annual technical reports) all clearly state that the SJT measures five of these attributes only (Coping with pressure, Working effectively as part of a team, Effective communication, Problem solving, and Commitment to professionalism). It is clearly acknowledged that the SJT would not be an appropriate methodology for assessing all nine attributes and this information is widely available. Importantly, the SJT in question was designed with clinicians and trainees using real scenarios encountered by trainees in clinical practice.

Najim et al suggest that there is a lack of standardization when assessing students ranked “middle” at the UK’s best and worst medical schools. Which do they perceive as the UK’s best and worst medical schools and on what evidence do they define them as such? The authors refer to Simon et al’s study that found no correlation between the Educational Performance Measure and the SJT. However, this preliminary study is based on a sample of only 239 applicants in one location. In larger-scale study analyses undertaken year on year across the full sample of applicants (n = 8,000), the correlation between SJT and Educational Performance Measure is consistently around $r = 0.30$ (2013; $r = 0.30$; 2014; $r = 0.30$; 2015; $r = 0.34$), which is to be expected. These results are made available publicly each year in an annual technical report published on the Foundation Programme website, and it would be prudent to also refer to full sample data as the basis for commentary. In recent studies, importantly the Foundation SJT positively correlates with subsequent in-training supervisor ratings and so provides good evidence of predictive validity.

The Foundation SJT is specifically designed to measure nonacademic attributes important for trainees entering clinical practice, and it is important that these attributes are measured reliably at the point of entry to the Foundation Programme in addition to a measure of academic performance. We agree with Najim et al in highlighting that examinations up until this point are largely based on academic performance from General Certificate of Secondary Education (GCSE) exams to medical school finals. Junior doctors go through a gruelling 9 years of back-to-back examinations, making them the UK’s most academically assessed students.

It is crucial to note that the possession of a medical degree has never been the only requirement for recruitment into the Foundation Programme. As with any job application, a requirement to consider applicants’ ability to meet the person specification is the most important requirement, possession of a degree being only a part of the person specification. Initially, “white space” questions were used in varying formats to determine this. However, it was agreed that this form of test had limitations, and therefore work was done to consider what would replace this, ultimately leading to the introduction of the SJT.

Najim et al conclude that the allocation of placements to junior doctors would be better served through an interview process. In this context, the use of interviews was reviewed, and while structured interviews can offer benefits as a selection method, they also have several limitations. Most prevalent are the resources required and the practicalities and costs associated with undertaking approximately 8,000 interviews, and ensuring that this is done in a robust and standardized manner. These issues, among others, were explored in depth as part of an Options Appraisal carried out in 2010, which assessed the relative benefits of the SJT alongside structured interviews as well as multiple mini interviews. The appraisal concluded that interviewing over 8,000 potential trainees was simply not feasible, would lack appropriate standardization, and would be too costly for a health service already severely constrained by resources.

In conclusion, the evidence widely available from both the UK Medical Schools Council and the Foundation Programme indicates that an SJT is an appropriate and fair measurement method to be used as one part of selection to the Foundation Programme, which targets important professional attributes required by all clinicians in practice.

**Disclosure**

Katie Petty-Saphon is the Chief Executive of the UK Medical Schools Council, which is funded by HEE to develop and deliver the Foundation Programme SJT. Kim Walker is affiliated with the UK Foundation Programme Office which oversees the delivery and scoring of the UK Foundation Programme SJT. Fiona Patterson, Vicki Ashworth, and Helena Edwards are affiliated with Work Psychology Group, which provides consulting advice and support to Health Education England on selection methodology. Work Psychology Group does not receive royalties for any methodology used. The authors report no other conflicts of interest in this work.

**References**


