

Could situational judgement tests be used for selection into dental foundation training?

F. Patterson,¹ V. Ashworth,² S. Mehra³ and H. Falcon⁴

VERIFIABLE CPD PAPER

IN BRIEF

- Questions the current method of selection into dental foundation training and its ability to meet the requirements of recruitment best practice.
- Demonstrates the overall benefits of a situational judgement test (SJT) for dental foundation training selection over current selection methods.
- Outlines plans to implement a SJT nationally for all candidates applying for dental training posts in 2012.

Objective To pilot and evaluate a machine-markable situational judgement test (SJT) designed to select candidates into UK dental foundation training. **Design** Single centre pilot study. **Setting** UK postgraduate deanery in 2010. **Participants and methods** Seventy-four candidates attending interview for dental foundation training in Oxford and Wessex Deaneries volunteered to complete the situational judgement test. **Interventions** The situational judgement test was developed to assess relevant professional attributes for dentistry (for example, empathy and integrity) in a machine-markable format. Test content was developed by subject matter experts working with experienced psychometricians. **Main outcome measures** Evaluation of psychometric properties of the pilot situational judgement test (for example, reliability, validity and fairness). Scores in the dental foundation training selection process (short-listing and interviews) were used to examine criterion-related validity. Candidates completed an evaluation questionnaire to examine candidate reactions and face validity of the new test. **Results** Forty-six candidates were female and 28 male; mean age was 23.5-years-old (range 22-32). Situational judgement test scores were normally distributed and the test showed good internal reliability when corrected for test length ($\alpha = 0.74$). Situational judgement test scores positively correlated with the management, leadership and professionalism interview ($N = 50$; $r = 0.43$, $p < 0.01$) but not with the clinical skills interview, providing initial evidence of criterion-related validity as the situational judgement test is designed to test non-cognitive professional attributes beyond clinical knowledge. Most candidates perceived the situational judgement test as relevant to dentistry, appropriate for their training level, and fair. **Conclusion** This initial pilot study suggests that a situational judgement test is an appropriate and innovative method to measure professional attributes (eg empathy and integrity) for selection into foundation training. Further research will explore the long-term predictive validity of the situational judgement test once candidates have entered training.

INTRODUCTION

Completion of dental vocational training (foundation year one) is a requirement for all UK dental graduates wishing to work as an NHS performer. The number of UK dental foundation places is matched to predicted numbers of graduates from UK dental schools; however, there is increasing competition for these places from overseas qualified dentists who are also eligible

to apply. Shortlisting of candidates via application forms is not only difficult and time consuming to conduct reliably when large numbers of applications are received, but it can be hard to prevent plagiarism with the application forms currently used. With the advent of nationally-coordinated selection in future, ensuring calibration and standardisation of selection methods will be challenging. A demonstrably robust system to rank applicants accurately is essential to ensure fair recruitment practice and to provide a robust defence against potential legal challenges.

Historically, approved training practices recruited trainees each year on an individual basis. However, following allegations of unfair recruitment practices, UK dental deaneries have increasingly taken on this responsibility, directly recruiting dental foundation trainees and assigning them to approved training practices.

The usual method of allocating trainees to training places has been on the basis of trainee preferences, prioritised according to merit based on interview performance. Sometimes there can be an additional element of matching, taking trainers' preferences into account following short meet-and-greet events. In 2009, Oxford and Wessex Deaneries investigated whether situational judgement tests (SJTs) could be used as a fair and robust way to assess important non-cognitive professional attributes as part of a managed recruitment process for dental foundation training posts.

SJTs have been used as a selection method in a variety of professions including the military, accountancy, the police and more recently in medicine.¹ There is not one single type of SJT – they are a measurement method designed to assess individuals' judgement regarding

¹University of Cambridge, Free School Lane, Cambridge and Work Psychology Group, Derbyshire; ²Work Psychology Group, Derbyshire; ³Associate Dental Dean, Oxford and Wessex Deaneries, the Triangle Roosevelt Drive, Headington, Oxford; ⁴Postgraduate Dental Dean, the Triangle Roosevelt Drive, Headington, Oxford and Vice Chair COPDEND

*Correspondence to: Professor Fiona Patterson
Email: fcp27@cam.ac.uk

Refereed Paper

Accepted 1 March 2012

DOI: 10.1038/sj.bdj.2012.560

©British Dental Journal 2012; 213: 23-26

situations encountered in the workplace. Candidates are presented with a set of hypothetical work-based scenarios and make judgements about possible responses, which are evaluated against a pre-determined scoring key,² enabling the SJT to be machine-marked (see Fig. 1 for an example item). Over the last 20 years, SJTs have become increasingly popular in large-scale selection, typically at the shortlisting stage.³ SJTs have been shown to have significant validity in predicting future job performance and can offer incremental (additional) validity over other selection methods.⁴⁻⁶ SJTs tend to produce smaller scoring differences between demographic groups than ability tests (eg based on ethnicity),⁷ have high face validity⁸ and are generally perceived positively by candidates.⁵

SJTs have been recently developed and implemented for postgraduate medical selection in the UK. For example, Patterson and colleagues⁹ developed a SJT for shortlisting candidates for GP training, which was implemented nationally in 2007. Research has demonstrated that the GP SJT is the best single predictor of performance at the final stage of UK GP selection and is a strong predictor of performance during and at the end of training licensing examinations.¹⁰ SJTs have since been developed and piloted to assess relevant professional attributes for other postgraduate specialties, such as core medical training, anaesthesia, surgery and public health.¹¹ An SJT has been successfully piloted for selection into foundation training for medicine in the UK, which will be launched for nationally-coordinated recruitment from 2013.¹²

This paper describes an initial 'proof of concept' study to examine whether a machine-markable SJT can provide a valid method for assessing professional attributes at selection into dental foundation training. To the authors' knowledge, this is the first known application of a SJT method to dentistry and will inform the debate regarding selection practices for postgraduate training.

With any new selection methodology, various psychometric and legal criteria must be satisfied including standardisation, reliability, validity and fairness.¹³ This paper reports on the development of the SJT and its psychometric properties when piloted with a sample of dental foundation

A new patient, Louise, asks you to veneer all her front teeth. She hopes to pursue a career as a model and has been advised by a friend that veneers may help her to do this. On examination, Louise's teeth are perfectly healthy with no previous restorations and just some very mild crowding. Her teeth are a vita shade A3.

Rank in order the following actions in response to this situation (1 = most appropriate; 5 = least appropriate).

- A. Respect Louise's wishes and schedule an appointment for the veneers.
- B. Ensure Louise is aware of alternatives such as orthodontics and tooth whitening.
- C. Suggest a course of tooth whitening and explain the legal issues involved.
- D. Establish exactly what Louise does not like about her teeth.
- E. Offer to refer Louise to a specialist to discuss the veneers in more detail.

Fig. 1 Example SJT item

Clinical interview:

A new patient presented with symptoms of reversible pulpitis and upon examination she has a deep cavity in a molar tooth that extends subgingivally. The patient wants a white filling. What further information would you need and what would you do?

Management, leadership and professionalism interview:

You have been working for three weeks in your training practice and are constantly running late. The nurse has complained to the practice manager. What approach would you take to resolve the situation?

Fig. 2 Example questions for clinical skills interview and the management, leadership and professionalism interview

candidates. SJT scores were also compared with the current selection process, comprising an application form and interview process. The objectives of this paper were to explore the reliability, validity and fairness of the SJT for dental foundation candidates. Specifically, we addressed the following research questions:

1. Are the psychometric properties of the SJT robust?
2. Is the SJT a valid method of selection for dental foundation trainees; that is, do the SJT scores correlate with performance in the interview process?
3. Does the SJT produce scoring differences based on demographic groups (ie gender, ethnicity)?
4. Do candidates react to the SJT positively and view it as fair; that is, does the SJT have good face validity?

METHOD

Design of the SJT

The SJT was designed specifically to assess professional attributes for dentistry in a machine-markable format to aid efficiency and standardisation. The principles of the test method were adapted from a validated test used for selection into UK general practice training.⁹ The SJT targets four key professional attributes from the dental foundation curriculum: professional integrity, teamworking, empathy and communication, and resilience and coping

with pressure. Test content was developed by a group of eight subject matter experts (experienced dental trainers and educators) working with experienced psychometricians and was designed to be specific to dentistry without requiring dental experience beyond that expected of all candidates. The SJT pilot paper contained 20 items to be completed in 30 minutes.

Procedure and sample

Data collection took place during the 2010 dental foundation training recruitment process at NHS South Central (Oxford and Wessex Deaneries). The selection process comprised an application form followed by a two-station structured interview process. The first interview targeted clinical skills, where candidates were presented with two clinical scenarios in ten minutes and asked how they would respond. Similarly, the second interview targeted management, leadership and professionalism, where candidates were presented with two scenarios in ten minutes relating to professional issues and asked how they would respond (see Fig. 2 for example interview questions). There were two interviewers for each station, who scored candidate responses independently on a scale of 1 to 5 (where 1 = poor and 5 = excellent) and scores were combined to give an overall score out of 20.

Of the 253 candidates who applied, 120 were shortlisted for interview. All

candidates attending the interview were invited to complete the SJT pilot for research purposes only. Seventy-four volunteer candidates participated after giving informed consent (62% response rate). Sixty-two percent were female and 38% male. Mean age was 23.5 years (range 22-32) and 60% of candidates described themselves as Asian, 23% White, 3% Black, 3% mixed, 3% Chinese and 9% other. Cronbach's alpha and the Spearman-Brown formula were used to evaluate the reliability of the SJT on all candidates sitting the test. Correlations between scores on the SJT and the two interview stations were analysed using Pearson's r , of which data were available for 50 candidates. The candidate evaluation questionnaire consisted of five items focusing on the relevance and fairness of the SJT using a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). Example items include '*the content of the situational judgement test was relevant to dental training*' and '*the content of the situational judgement test appeared to be fair*'.

RESULTS

Are the psychometric properties of the situational judgement test robust?

SJT total scores for the candidate show approximately normal distributions, indicating that the test was capable of differentiating between individuals, which is important in a selection context for ranking purposes. Results show the reliability of the SJT to be $\alpha = 0.74$, which demonstrates sufficient reliability for a pilot test.⁹

Does the SJT predict performance in the shortlisting and interview process?

Shortlisting and interview scores were normally distributed as was the distribution of SJT scores. Correlations (Pearson's r) between pilot SJT scores, shortlisting scores and live interview scores were calculated and are presented in Table 1. There was no correlation between shortlisting scores and the SJT, which is to be expected as shortlisting focuses on previous qualifications and experience whereas the SJT is a written test, focusing in non-cognitive attributes. Results show

Table 1 Descriptive statistics for the interview stations and correlations with the SJT

	Shortlisting total score	Interviews overall score	Clinical skills interview	Management, leadership and professionalism interview
N	49	50	50	50
Mean score	64.2	31.3	15.2	16.1
Std dev	4.9	4.7	3.9	2.1
Range	48-73	18-40	6-20	10-20
Correlation with SJT total score	$r = 0.21$	$r = 0.34^*$	$r = 0.18$	$r = 0.43^{**}$

Key: * $p < 0.05$, ** $p < 0.01$

a moderately positive correlation between the SJT scores and the overall interview score ($r = 0.17$, $p < 0.05$), which is to be expected as the interviews were based on work relevant scenarios. Specifically, results show differential results for each of the two interview stations. There was no significant correlation between the SJT scores and the clinical skills interview; however, the correlation between the SJT and the management, leadership and professionalism interview shows a strong significant positive correlation ($r = 0.43$, $p < 0.01$). In combination, these results provide good evidence of criterion-related validity of the SJT as it is designed to assess professional non-cognitive attributes, which are also targeted in the management, leadership and professionalism interview station as opposed to clinical knowledge measured in the clinical skills interview.

Does the SJT produce scoring differences based on demographic groups?

Group differences in performance were analysed on the basis of gender, age and ethnic group. Female candidates scored slightly but significantly higher than male candidates (by approximately 0.5 SD; t -test $p < 0.05$); this is a typical finding in selection across many occupational groups.¹⁴ There was no significant correlation (Spearman's rho) between age and scores on the SJT. In terms of ethnic group, a t -test showed that there was no significant difference in performance between White and Asian candidates (sample sizes were too small to compare other groups). Overall, few group differences were found in performance on the SJT, which is reassuring in relation to fairness issues.

Do candidates perceive the SJT to be fair and relevant?

All 74 candidates completed an evaluation questionnaire to examine the face validity of the SJT. Eighty-nine percent of candidates agreed or strongly agreed that the SJT was relevant to dental training, 85% thought it was appropriate for their training level and 78% felt that it was fair. These results demonstrate good evidence for the face validity of the SJT and are commensurate with findings elsewhere when evaluating SJTs for use in selection.⁹

DISCUSSION

This proof of concept study provides initial evidence of the effectiveness of a SJT method to assess professional attributes at selection into UK dental foundation training, representing a significant innovation in the selection process. Professional attributes are clearly important for any trainee to demonstrate when entering a career in dentistry and these are also assessed during training. Results suggest that the SJT approach shows potential as a selection tool for dental foundation training, although further work on item development and piloting is needed to ensure an effective test for live selection. The SJT showed good internal reliability when corrected for test length and was capable of ranking candidates fairly. There were few group differences in performance on the SJT, although this aspect should be monitored in any future research. There was no significant correlation between scores on the SJT and in the shortlisting, which is to be expected as shortlisting focuses on prior qualifications and experience and the SJT focuses on aptitude relating to key professional attributes. In addition, since the SJT correlates significantly with

performance at interview, especially in relation to performance in management, leadership and professionalism scenarios, results indicate good initial validity of the SJT. Note that the size of the correlation is not very strong ($r = 0.70+$), which indicates the SJT provides additional information about the aptitude of candidates over the interviews. Candidate reactions to the pilot SJT were very positive, with most rating the SJT as relevant to dentistry, appropriate for the target training level and fair.

There were some limitations of this research that should be noted. Firstly, the sample size for this initial pilot was small and results should therefore be interpreted with caution. However, despite the small sample size, this proof of concept study shows encouraging initial findings. Secondly, results of the item analysis indicated that further work is needed in order to develop a bank of effective items. A further follow-up study is underway, commissioned by the Committee of Postgraduate Dental Deans and Directors (COPDEND) using a larger sample of applicants to 2011 training posts in a number of English deaneries. In 2012, recruitment will be carried out on a national basis, coordinated in England and Wales, with deaneries in Scotland and Northern Ireland operating separate parallel processes. In this respect it will be yet more important to ensure there is a fair, robust and objective scoring system that allows ranking

of candidates to comply with a meritorious system of selection and allocation to training places. Machine-markable tests, such as SJTs, provide standardised selection tools that can complement other selection methods in building a fuller picture of a candidate's skills. They are also likely to increase the utility of the selection process once the initial development phase is complete. Best practice would dictate that SJTs would be used alongside interviews as part of a suite of methods with the aim of ensuring an efficient and effective selection method.

Professor Simon Plint is acknowledged for his contribution in the original study design. Victoria Carr is acknowledged for her significant contribution in data collection, preparation of data sets, reporting and data analysis. Branwen Thomas is acknowledged for her significant contribution to support in data collection and organisation. The various contributors from the locations helping in data access are acknowledged. Source of funding is UK Committee of Postgraduate Dental Education Directors (COPDEND).

Competing interests

F. Patterson and V. Ashworth provide advice to the Department of Health in the UK on selection and recruitment issues through the Work Psychology Group.

Authorship and contributorship

F. Patterson conceived of the original study and design, analysed and interpreted the data, and wrote the paper. V. Ashworth contributed to the overall study design, analysed and interpreted the data and helped write the paper. H. Falcon and S. Mehra conceived of the original study and contributed to the overall study design, helped in data collection, interpreted the data and helped write the paper. All authors commented on the final version of the paper.

1. Irish B, Carr A, Sowden D, Douglas N, Patterson F. Recruitment into specialty training in the UK. *BMJ Careers*, 2011.
2. Weekley J A, Ployharty R E (eds). *Situational judgment tests: theory, measurement and application*. Society for Industrial and Organizational Psychology, 2006.
3. Lievens F, Peeters H, Schollaert E. Situational judgment tests: a review of recent research. *Pers Rev* 2008; **37**: 426–441.
4. McDaniel M A, Hartman N S, Whetzel D L, Grubb III W L. Situational judgment tests, response instructions and validity: a meta-analysis. *Pers Psychol* 2007; **60**: 63–91.
5. Chan D, Schmitt N. Situational judgment and job performance. *Hum Perform* 2002; **15**: 233–254.
6. Lievens F, Patterson F. The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *J Appl Psychol* 2011; **96**: 927–940.
7. Lievens F, Buyse T, Sackett P R. The operational validity of a video-based situational judgment test for medical college admissions: illustrating the importance of matching predictor and criterion construct domains. *J Appl Psychol* 2005; **90**: 442–452.
8. Clevenger J, Pereira G M, Wiechmann D, Schmitt N, Harvey V S. Incremental validity of situational judgment tests. *J Appl Psychol* 2001; **86**: 410–417.
9. Patterson F, Baron H, Carr V, Plint S, Lane P. Evaluation of three short-listing methodologies for selection into postgraduate training in general practice. *Med Educ* 2009; **43**: 50–57.
10. Patterson F, Lievens F, Kerrin M, Munro N, Irish B. The predictive validity of selection for entry into postgraduate training. Under review, *Med Educ* 2011.
11. Patterson F, Carr V, Zibarras L *et al*. New machine-marked tests for selection into core medical training: evidence from two validation studies. *Clin Med* 2009; **9**: 417–420.
12. *Improving selection to the foundation programme*. UK: Medical Schools Council, 2012. Online information available at <http://www.isfp.org.uk/Pages/default.aspx> (last accessed June 2012).
13. Schmidt F L, Hunter J E. The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychol Bull* 1998; **124**: 262–274.
14. Patterson F, Denney M-L, Wakeford R, Good D. Fair and equal assessment in postgraduate training? A future research agenda. *Br J Gen Pract* 2011; **61**: 712–713.