

Values Based Recruitment



Guidance for using Situational Judgement Tests for Values Based Recruitment

Developing people
for health and
healthcare

www.hee.nhs.uk

Acknowledgements

Health Education England (HEE) would like to thank and acknowledge Work Psychology Group for their support when completing 'Guidance for using Situational Judgement Tests for Values Based Recruitment'.

Work Psychology Group has over 20 years' experience of designing and evaluating assessment and recruitment methodologies for high stakes selection in healthcare, finance, engineering and professional services. The team is made up of experienced consultants who have worked in both research and industry, giving a unique blend of organisational and academic expertise, ensuring the latest research can be turned into practical, innovative solutions. For more information, please see www.workpsychologygroup.com.

Contents

Section	Page
Overview	4
Summary	5
What is a Situational Judgement Test?	6
What is the evidence base for SJTs?	9
What makes an SJT values based?	10
Key considerations for implementing an SJT for VBR	10
<ul style="list-style-type: none"> Developing a test 	10
<ul style="list-style-type: none"> Test administration and security 	11
<ul style="list-style-type: none"> Applicant perceptions 	12
<ul style="list-style-type: none"> Integrating with other selection methods 	12
<ul style="list-style-type: none"> Feedback 	12
Example case studies	13

Overview

The following guidance is designed to highlight key considerations NHS employing organisations and admissions staff within Higher Education Institutions (HEIs) may need to take into account implementing a Situational Judgement Test (SJT). It is intended to be used as a guide to assist with the planning, development and implementation of selection activities whilst recognising the need for individuals to tailor their approach to meet the needs of their own organisation.

It is recommended that utilisation of recruitment methods for assessment of values need to be considered with regards to the national core requirements, as set out in the national VBR framework. This recommends that an individual structured interview (including MMIs) to assess values must be conducted as part of the overall recruitment process to effectively recruit for values. This may occur in isolation or as part of a wider process such as a selection centre or in conjunction with a screening method such as an SJT.

See the [HEE website](#) for more information about the national core requirements and to learn more about structured interviews.

Summary

- There is evidence that SJTs show improved validity, favourable applicant reactions and smaller sub-group differences than other selection methods.
- Since SJTs are machine-markable and the scoring keys are predetermined they offer significant advantages over other methods that are hand-scored by assessors (such as in selection centre exercises or personal statements) including reduced assessor time and reduced assessor bias.
- SJTs can be designed to measure a variety of non-technical attributes, in a healthcare setting they can be designed to assess non-clinical attributes (such as empathy and integrity) which can be mapped to organisational values.
- SJTs which are developed as a bespoke tool and are based on thorough role analysis, and involve Subject Matter Experts (SMEs) are likely to have improved levels of reliability, validity and fairness.
- SJTs which are delivered to a high volume of applicants and/or in an unproctored setting are most at risk to items becoming over-exposed. Therefore, for these tests in particular it is important to ensure processes are in place for on-going and regular refreshing of items or development of new content.
- The development of bespoke SJTs can be relatively lengthy and resource-intensive, and therefore this approach is likely to be most appropriate for high volume selection processes. Coordinated efforts (for example, across a consortium of a number of organisations) to develop, pilot and continue to refresh/refine content are therefore most likely to result in high quality tests based on robust data analysis.

What is a Situational Judgement Test?

In a Situational Judgement Test (SJT), test-takers are presented with written or video-based depictions of hypothetical scenarios and asked to identify an appropriate response from a list of alternatives. A variety of response formats can be used and these are typically classified into one of two formats: **knowledge based** (i.e. what is the best option) or **behavioural tendency** (i.e. what are you most likely to do). SJTs are typically scored by comparing applicants' responses to predetermined scoring keys agreed by subject matter experts (SMEs).

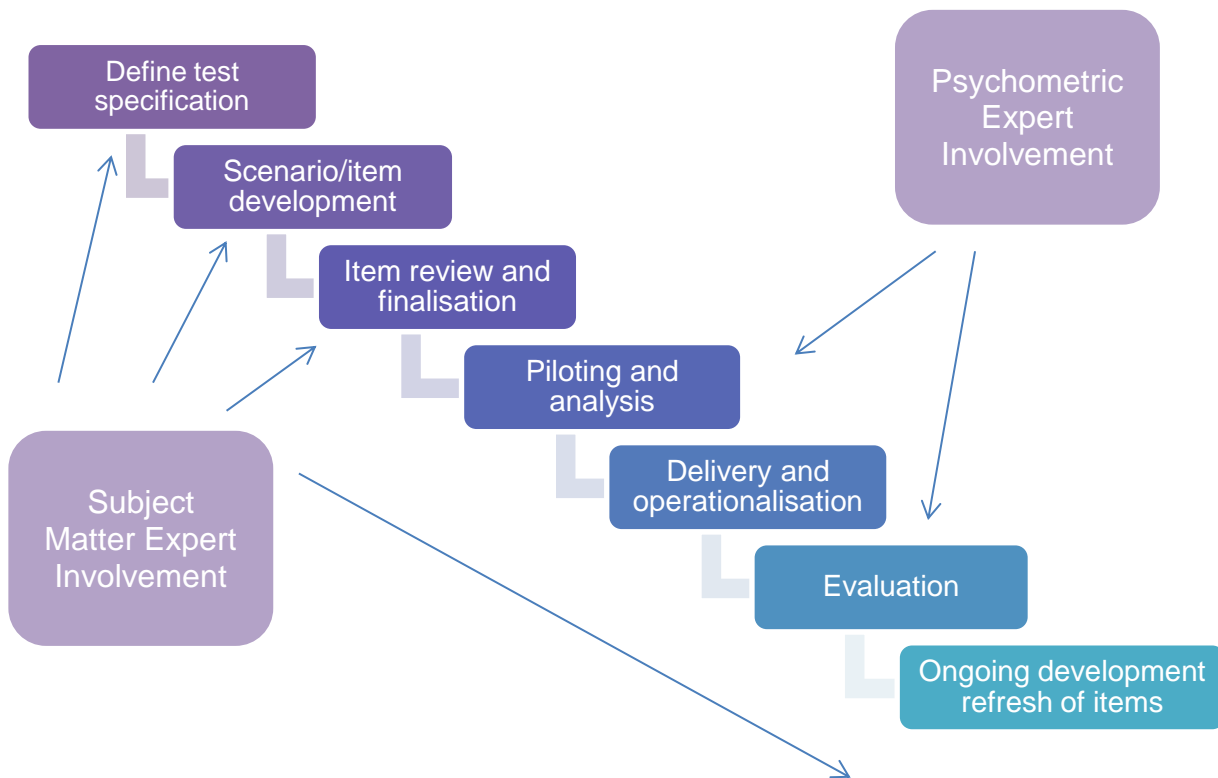
SJTs are different to other selection methods which present applicants with hypothetical situations (such as situational interviews) as they have predetermined, fixed response choices and can be machine-markable. In this respect SJTs can offer significant advantages for high volume assessment because they are less resource intensive to administer and score. They also have the benefit of being less subject to assessor bias than other face-to-face forms of selection which rely on assessor/interviewer judgement.

SJTs are usually developed as a bespoke tool for a particular role or group of roles. This development process should be based on a thorough role analysis which identifies criteria which are relevant to successful performance in the role. The development of test items is then based around the output of the role analysis and typically involves SMEs generating ideas for relevant and challenging scenarios that are likely to be faced in the target role. Through consultation with SMEs, a series of possible responses are developed for each scenario which range in terms of their appropriateness or effectiveness. The involvement of SMEs helps to ensure that the scenarios are realistic and are seen as appropriate and relevant by the test takers. Therefore SMEs will ideally be involved throughout test development; particularly at the early stages where the test specification is being designed and items are being developed and refined.

Through a process of review and piloting, statistical analysis can be undertaken to identify items which are most likely to be reliable and valid measures of future performance.

On-going evaluation activity should also be in place to ensure that the test performance can be monitored over time. Ideally this evaluation activity would be supported by individuals with particular expertise in psychometric analysis. A diagram summarising the steps involved in developing a best practice SJT can be seen in Figure 1 overleaf, which highlights the stages where SME and psychometric expertise is particularly beneficial. Given the lengthy and resource-intensive nature of SJT design, this selection method is often used for high volume applications and developed through collaborative efforts across a number of organisations or through a consortium who can leverage access to a wide variety of stakeholders.

Figure 1 Overview of best practice SJT development process



Three example SJT items are presented in Figure 2, 3 and 4 overleaf. The example in Figure 2 is taken from the SJT element of the UK Clinical Aptitude Test (UKCAT) which is used for selection into medical school and the examples in Figures 3 and 4 are taken from the SJT used as part of the selection process into Foundation Year One (FY1) training for doctors. These examples illustrate some of the different approaches which can be taken in presenting SJT items.

In Figure 2 the scenario is presented in the third person and the test taker is required to rate the appropriateness of each of the four responses independently from one another (i.e. two responses could be rated as highly appropriate).

Figure 2 Example SJT item for selection to Medical School (Rating)

A consultation is taking place between a senior doctor and a patient; a medical student is observing. The senior doctor tells the patient that he requires some blood tests to rule out a terminal disease. The senior doctor is called away urgently, leaving the medical student alone with the patient. The patient tells the student that he is worried he is going to die and asks the student what the blood tests will show.

How appropriate are each of the following responses by the medical student in this situation?

(1= very appropriate, 2 = appropriate but not ideal, 3 = inappropriate but not awful, 4 = very inappropriate)

- A. Explain to the patient that he is unable to comment on what the tests will show as he is a medical student.
- B. Acknowledge the patient's concerns and ask whether he would like them to be raised with the senior doctor.
- C. Suggest to the patient that he poses these questions to the senior doctor when he returns.
- D. Tell the patient that he should not worry and that it is unlikely that he will die.

In Figure 3, the scenario is presented in the second person and the test taker is required to select three out of the eight possible responses which would most appropriate address the situation.

Figure 3 Example SJT item for selection to the UK Foundation Training Programme (Multiple Choice)

You review a patient on the surgical ward who has had an appendicectomy done earlier on the day. You write a prescription for strong painkillers. The staff nurse challenges your decision and refuses to give the medication to the patient.

Choose the THREE most appropriate actions to take in this situation

- A. Instruct the nurse to give the medication to the patient
- B. Discuss with the nurse why she disagrees with the prescription
- C. Ask a senior colleague for advice
- D. Complete a clinical incident form
- E. Cancel the prescription on the nurse's advice
- F. Arrange to speak to the nurse later to discuss your working relationship
- G. Write in the medical notes that the nurse has declined to give the medication
- H. Review the case again

In Figure 4, the scenario is also presented in the second person but this time five responses are presented which are to be ranked in order from the most to least appropriate response.

Figure 4 Example SJT item for selection to the UK Foundation Training Programme (Ranking)

You are reviewing a routine drug chart for a patient with rheumatoid arthritis during an overnight shift. You notice that your consultant has inappropriately prescribed methotrexate 7.5mg daily instead of weekly.

Rank in order the following actions in response to this situation (1= Most appropriate; 5= Least appropriate)

- A. Ask the nurses if the consultant has made any other drug errors recently.
- B. Correct the prescription to 7.5mg weekly.
- C. Leave the prescription unchanged until the consultant ward round the following morning.
- D. Phone the consultant at home to ask about changing the prescription.
- E. Inform the patient of the error.

As can be observed through examining the three examples above, SJTs can be pitched at different levels depending on the target role. For example, for items which are used as part of selection on to the Foundation Programme a degree of medical knowledge is assumed as all applicants will have completed medical training. Such medical knowledge is not assumed for items which are designed to be completed at the point of entry to Medical School (i.e. the UKCAT example). The response format should be chosen based on the requirements of the target role and the population which are to be assessed.

What is the evidence base for SJTs?

Although the SJT methodology has been in existence for several decades, in the past 10 years their use has become increasingly popular in large-scale selection across all occupational groups internationally.

There is good evidence to show that appropriately designed SJTs are a useful methodology to evaluate a range of professional attributes for selection into medicine and dentistry, for both admissions (for example, the UKCAT SJT) and for entry into postgraduate training (for example, the FY1 SJT and Dental foundation SJT). Long-term follow-up studies have shown an SJT measuring empathy, integrity and resilience (used to select applicants applying for training in UK General Practice) to be the best single predictor of subsequent job performance

and licensing outcomes compared to other selection methods^{1 2}. SJTs have also been shown to be less susceptible to demographic group differences than other selection tools³. For a more in-depth summary of the evidence-base relating to SJTs please see the **VBR literature review** on the [HEE website](#).

What makes an SJT values based?

SJTs can be designed to assess values which are important to the role as part of the domains targeted in a test specification. However, not all SJTs are specifically values based and some may be designed to measure other constructs such as specific job (or procedural) knowledge.

The theory underlying SJTs suggests they measure implicit trait policies (ITPs) as well as general experience (and potentially job-specific knowledge depending on the role). ITPs are beliefs about the relative costs and benefits of expressing particular traits (which guide behaviour) in certain situations. For example, making a judgement that being agreeable in a situation is likely to be a more successful strategy than being disagreeable. In this way, SJTs measure an individual's awareness about what is effective behaviour in a given situation and this is likely to be linked to an individual's values.

Like values, it is thought that ITPs are shaped by general socialisation and life experiences such as parental modelling during childhood that, for example, teach the value of certain agreeable expressions such as: helping others in need, turning the other cheek or looking after one's neighbours; or disagreeable expressions such as: showing selfish preoccupation with one's own interests, holding a grudge/'getting even' or advancing ones own interests at another person's expense. Therefore, it is suggested that since SJTs measure ITPs, they are well placed to provide insight into aspects of an individual's values.

Key considerations for implementing an SJT for VBR

Developing a test

- Most SJTs are developed as a bespoke tool so that they can be tailored to the specific job role, as this helps to create scenarios which are relevant and realistic. Such tests tend to be viewed more positively by applicants who can see the direct relevance of the questions asked to the job for which they are applying. Furthermore, such a bespoke approach has the advantage of presenting applicants with a realistic job preview, providing an insight into the day-to-day challenges which might be experienced in the role. This gives SJTs the potential to serve a useful purpose as a self-selection tool. The involvement of service users in addition to SMEs in the test

¹ L Lievens F and Patterson F. (2011). 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. *Journal of Applied Psychology*, 96(5), 927-940.

² Patterson F, Lievens F, Kerrin M, Munro N and Irish B. (2013). The predictive validity of selection for entry into postgraduate training in general practice: evidence from three longitudinal studies *British Journal of General Practice* 11; 63(616):734-741

³ Clevenger J, Pereira GM, Wiechmann D, Schmitt N and Harvey VS. (2001). Incremental validity of Situational Judgment Tests. *Journal of Applied Psychology*, 86(3):410-417.

development can also be valuable.

- Whilst enhancing the relevance and realism of the scenarios, developing bespoke SJTs through the involvement of SMEs can be resource intensive.
- As with any psychometric instrument used to make or inform selection/assessment decisions, an SJT should be subject to analysis which assesses the reliability, validity and fairness of the tool. This is usually achieved through thorough piloting, and ideally includes an assessment of the extent to which performance of the tool is predictive of performance in the job role (predictive validity). Through analysis of pilot data, it is possible to establish appropriate cut-scores to determine how test scores are used to make selection decisions.
 - It is important that adequate pilot data are gathered to ensure robust analysis can be undertaken. For this reason, piloting ideally takes place with large samples of applicants which may necessitate the involvement of multiple organisations or to be undertaken over a lengthy period of time.
- To ensure the content of the test remains relevant and up-to-date, it is good practice to undertake an on-going process for developing and piloting new items. The regularity of such on-going development is likely to be determined in part by the way in which the test is administered and the number of applicants it is completed by (see test administration and security below).
- Given that the development of bespoke SJTs can be relatively lengthy and resource-intensive, this approach is likely to be most appropriate for high volume selection processes. Coordinated efforts, for example across a consortium of a number of organisations, to develop, pilot and continue to refresh/refine content are therefore most likely to result in high quality tests based on robust data analysis.

Test administration and security

- How and where the SJT will be completed is an important consideration. Ordinarily an SJT will be completed under timed conditions and therefore provision would need to be made for this, for example, within a selection centre timetable.
- If the test is completed in an unproctored (or unsupervised) setting (online, in the applicant's own time prior to a selection event), then there is a greater risk that the applicant may attempt to get help in responding to the items. There may also be a risk of another individual completing the test on their behalf, or that they may attempt to make a copy of the test questions to share with others (for example, taking computer screen shots to post in an online forum).
 - There are methods for guarding against such issues, for example, including a verification test later in the selection process against which an applicant's response in an unproctored setting are compared. Applicants can also be asked to sign non-disclosure agreements as a deterrent to sharing test

content. It is also possible to create multiple test versions which are subject to statistical equating processes to ensure equivalence across multiple test papers. This means that applicants can be presented with different test items reducing the likelihood that content will be shared.

- Fully proctored test administration is the most effective way of protecting test content and preventing cheating, however this can have practical and cost implications as it involves applicants attending a physical test centre. In situations where unproctored settings are a necessity, it is particularly important that test content is refreshed on a regular basis to avoid items becoming leaked.

Applicant perceptions

- The evidence shows that applicants generally react favourably to SJTs compared to other selection methods as they have higher levels of face validity and the relevance to the role is very clear (i.e. the questions asked make sense in relation to the role being applied for).
- However, it is still important that applicants are given a thorough explanation as to what the SJT involves and how the results will be used to inform the selection decisions. Where possible, it is also good practice to give applicants access to practice questions to enable them to understand the test format in advance.

Integrating with other selection methods

- If an SJT is used alongside other selection methods it is important to consider how the SJT score will contribute to the overall decision making process, for example what weighting will the SJT exert?
 - The weighting within the selection process should be agreed in advance and should be considered based on the person specification and role requirements.
- An alternative approach to integrating SJTs into the selection process is to use responses to particular SJT questions to inform more focused questions at interviewing. For example, an applicant could be asked to expand on why they responded in a particular way or what factors they took into consideration.

Feedback

- When providing feedback to applicants about their performance on the SJT, it is important to balance the provision of useful and meaningful feedback to the applicant while also protecting the content of the test. It is generally difficult to provide detailed feedback without revealing the details of the scoring key of particular items and thus compromising the security of the test items.
 - Meaningful feedback can however be provided which compares the applicant's

overall SJT score with others within the target population. Guidance can also be provided relating to the extent to which the applicant was aligned with SMEs. For example, those with higher scores tend to have responded in a way which is more consistent with SME judgement. When SJTs have been developed based on thorough role analysis information can also be provided within feedback which explains the domains which the SJT was designed to measure.

Example case studies

The [UK Clinical Aptitude Test \(UKCAT\)](#) is a bank of tests used by a consortium of UK university Medical and Dental schools as part of the admissions process to select applicants onto **Medical** and **Dental undergraduate degree programmes**. Since 2013, it includes an SJT which is designed to assess important non-cognitive attributes that are associated with the effective performance of medical and dental students: integrity, perspective-taking and team involvement. The development process involves an annual cycle of item development, review and piloting. Early evidence indicates that the UKCAT SJT is reliable and content valid, is differentiating effectively, and is supportive of widening participation. Work is planned to explore predictive validity at the soonest opportunity.

The [Foundation Year One \(FY1\) SJT](#) is used as part of the selection onto the **UK Foundation Programme** alongside a measure of educational attainment, each year assessing around 8000 applicants to determine allocation of Foundation School places. The SJT is designed to measure; commitment to professionalism, coping with pressure, effective communication, patient focus and working effectively as part of a team. These elements have been mapped to the values of the NHS Constitution. The ongoing development and piloting of the SJT is coordinated by the Medical Schools Council and involves a wide range of clinicians including Foundation Year 2 doctors. A predictive validation study is currently underway to explore the relationship between SJT scores and performance in-training.

[London South Bank University](#) is using an SJT as part of the screening process for the **Children's Nursing** programme. The development of the test involved input from a variety of children's nursing to help ensure the content was specifically relevant to the role of a children's nurse. On-going evaluation of the tool is in progress to monitor it's suitability to be used in this context. Initial evaluations demonstrate that the tool has helped to reduce the number of applicants invited to interview whilst achieving positive feedback from applicants and less discrimination against minority groups than other selection processes.

[University College Hospitals London Hospitals NHS Foundation Trust](#) uses a bespoke SJT which is completed by **all individuals** prior to making an application. This is used across all roles recruited for within the organisation. The SJT was developed through collaboration with existing staff to ensure the items are relevant and provide a tangible message about which values are important from the outset of the recruitment process. This is one part of broader initiatives to embed values across the organisation.