Developing people for health and healthcare

Evaluation of Values Based Recruitment (VBR) in the NHS

Analysis of VBR Activity within NHS Trusts

September 2014

THE NHS CONSTITUTION
the NHS belongs to us all

NHS
Health Education England

Acknowledgements

Contents

Appendices

1. Executive Summary

1.1. Values Based Recruitment (VBR) has been identified as a core objective in the NHS Health Education England (HEE) Mandate (April 2013 to March 2015). VBR is recognised as a key priority for HEE, which aims to deliver a system to recruit for values which promotes quality care and positive patient experience for NHS funded training posts and all new NHS employees by March 2015.

1.2. In order to support NHS Trusts and employing organisations to recruit for values, NHS Employers and HEE have partnered to develop an evidence-based framework with associated guidance to support organisations when choosing and implementing evidence-based selection methods within their local VBR programme.

1.3. To support the development of an evidence-based framework, a systematic review of the selection methods used by a sample of NHS Trusts to select for values was conducted. The aims of the review are as follows:

- Review a broad sample of commonly used selection methods to recruit for values;
- Evaluate the extent to which these sampled methods are in alignment with best practice using the evidence in the literature review[1];
- Use the experience of the sampled methods from Trusts to generate a series of lessons learned and issues for consideration, to serve as a useful reference point for other Trusts considering their own approaches to VBR; and
- Ensure that the real-life, practical experiences of Trusts currently engaged in VBR activity inform the development of the national VBR framework.

1.4. A sample of nine Trusts, identified by HEE and NHS Employers is included in this report. These Trusts were selected in order to provide a good spread of different evidence-based VBR methods (structured interviews, Situational Judgement Tests and selection centres), a broad coverage across professional roles assessed by the identified methods, and a diverse geographical coverage.

1.5. Information about the sampled VBR selection methods was collected from two sources: 1) Telephone interviews with representatives from identified Trusts; 2) Documentation in relation to the selection methods that are being utilised.

1.6. A summary of the results derived from the experiences of this cohort is presented below.

1.7 **Structured Interviews**

- Trusts using structured interviews generally see this as an effective method of improving recruitment decisions however they are also seen as resource-intensive.

---

[1] http://hee.nhs.uk/work-programmes/values-based-recruitment/vbr-evidence/

- The effectiveness of this method is reliant on the manner in which the interview is conducted and scored.
- There was evidence to support that a structured interview, when designed and implemented according to best practice, can predict subsequent role performance. However Trusts differed significantly in the degree of structure that is applied to the way in which the interviews are conducted.
- Organisational engagement and stakeholder consultation are seen as important aspects when developing structured interviews.
- Access to appropriate assessor/interviewer training is also an important consideration.

1.8 Situational Judgement Tests (SJTs)

- There is evidence of Trusts using SJTs; however these tend to be used at the screening stage to help shortlist suitable candidates.
- Developing and implementing an SJT according to best practice is time consuming and can be resource intensive but can provide a useful screening mechanism especially for roles with high volumes of applicants when appropriate and evidence based cut off scores are piloted and developed
- Practical constraints result in challenges with undertaking piloting and evaluation activity, which impacts on the value of the tools introduced.
- Tools which are designed to cover a broad range of roles and which have been developed without stakeholder consultation are likely to be less well received.
- The SJTs reviewed in this study were shown to be ineffective in screening out large numbers of applicants. In order to develop robust and fit-for-purpose SJTs there is a requirement for investment in on-going development and evaluation. This might be achieved through Trusts pooling resources and taking a consortium approach to the development of such tools.

1.9 Selection centres (SC)

- There is evidence that Trusts have achieved improved recruitment outcomes following the introduction of an SC however the resource-intensive nature of the implementation is a key consideration, and for this reason they may be considered appropriate only for certain roles.
- The overall effectiveness of an SC is reliant on the extent to which assessors are able to extract and score behavioural evidence. It is therefore important that assessors are given adequate training.
- Despite the labour intensive nature of SCs, stakeholders had generally responded favourably to their introduction as they were seen as adding significant value to a selection process.

1.10 **Recommendations**

Recommendations for the VBR Programme

- **Recommendation 1a:** Generate practical case study material, guidance documents, resources, training and support to be shared with employers to coincide with the launch of the VBR framework in October 2014.
- **Recommendation 1b:** Design mechanisms to further enhance understanding of NHS Constitution values within Trusts to help ensure their more consistent promotion and application, and provide resources, such as the NHS Employers Values Mapping tool, to enable organisations to map their own local values with those of the NHS Constitution.
- **Recommendation 1c:** Coordinate resources for Trusts to access to support the development, implementation and evaluation of VBR

Recommendations for individual NHS Trusts when undertaking VBR

- **Recommendation 2a:** Conduct a job/role analysis to systematically identify the values to be assessed, in relation to the NHS Constitution.
- **Recommendation 2b:** Incorporate stakeholder consultation when developing the selection method to maximise organisational engagement and to ensure a high degree of relevance to the role.
- **Recommendation 2c:** Provide training to assessors to ensure appropriate use of the selection method and ensure that evidence of a candidate's values may be extracted and scored in a standardised way.
- **Recommendation 2d:** Implement on-going evaluation measures to ensure continued improvements in selection method(s) including monitoring of possible sub-group differences.

2. Introduction

Background

2.1.   This report captures and reviews information about Values Based Recruitment (VBR) activity based on a case-study exercise which targeted a sample of NHS Trusts currently using VBR processes to select employees. The purpose of this report is to detail the results from the review and draw out key conclusions and recommendations to inform the development of the national VBR framework.

2.2.   VBR is an approach which attracts and selects students, trainees or employees on the basis that their individual values and behaviours align with the values of the NHS Constitution alongside skills and aptitude. This will ensure that the future and current NHS workforce is selected against the values of the NHS Constitution, to build a workforce not only with the right skills and in the right numbers, but with the right values to support effective team working in delivering excellent patient care and experience.

2.3.   VBR can be delivered in a number of ways; through screening assessments, to values based interviewing techniques, role play, written responses to scenarios, and assessment centre approaches.

2.4.   This report complements two separate reports; 1) a literature review[2] which presents the evidence base on the selection methods available to assess values and 2) an analysis of VBR activity within Higher Education Institutions (HEIs) offering NHS funded pre-registration healthcare programmes in England[3]. Taken together, these reports ensure that the VBR programme is supported by an evidence-based approach, incorporating findings both from the research literature and the experiences of HEIs and Trusts undertaking recruitment activity in practice.

Drivers for Values Based Recruitment

2.5.   When Health Education England (HEE) was established on 1 April 2013, the following was included in the Health Education England Directions 2013: *"The Secretary of State directs that HEE must exercise its functions under the HEE Directions with a view to ensuring that education and training for healthcare workers is provided in a way which promotes the NHS Constitution"* (paragraph 2.4). The NHS Constitution establishes the principles and values upon which NHS organisations should exist namely: Working together for patients; Respect and dignity; Commitment to quality of care; Compassion; Improving lives; and Everyone counts.

2.6.   The VBR programme is a priority for HEE. VBR was identified as a key deliverable in the 2013-2015 Mandate from the Government to HEE: "*HEE will oversee delivery of a national values based recruitment framework and associated tools and resources by October 2014 and ensure that selection into all new NHS funded training posts*

---

[2] http://hee.nhs.uk/work-programmes/values-based-recruitment/vbr-evidence/
[3] http://hee.nhs.uk/work-programmes/values-based-recruitment/vbr-evidence/

incorporates testing of values based recruitment by March 2015" (page 25). The focus on the NHS Constitution values across the NHS has been driven in part by the report of the Mid Staffordshire NHS Foundation Trust Public inquiry (Francis, 2013[4]) which highlighted the critical role that the workforce plays in ensuring the provision of high quality and safe health services and, in particular, the significance of staff values and behaviours on the level of care and patient experience. Other key drivers include reports such as Keogh (2013[5]), Berwick (2013[6]) and Cavendish (2013[7]) which have also served to place the spot-light on quality of care, patient safety, values, attitudes and aptitude for caring.

2.7. Although there is no evidence nationally, nor any suggestion from the above that recruitment strategies have failed to select people with the right values in the past, there is clearly a need to promote good recruitment practice and ensure that effective, evidence-based approaches to assess for values are put in place.

Objectives

2.8. As employers, NHS Trusts are responsible for determining approaches to recruitment locally. Therefore the focus of this study was to identify common issues in relation to VBR to enable the sharing of lessons learned and good practice. This study took a case-study approach whereby a number of NHS Trusts undertaking a range of VBR activity were sampled to share examples of their experiences of implementing VBR.

2.9. The objectives for this study are summarised below:

- Review a broad sample of commonly used selection methods to recruit for values;
- Evaluate the extent to which these sampled methods are in alignment with best practice using the evidence in the recent literature review;
- Use the experience of the sampled methods from Trusts to generate a series of lessons learned and issues for consideration to serve as a useful reference point for other Trusts considering their own approaches to VBR; and
- Ensure that the real-life, practical experiences of Trusts currently engaged in VBR activity inform the development of the national VBR framework.

---

[4] Francis, R. (2013). Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry.
[5] Keogh, B. (2013). Review into the quality of care and treatment provided by 14 hospital Trusts in England: overview report.
[6] Berwick, D. (2013). A promise to learn – a commitment to act: improving the safety of patients in England.
[7] Cavendish, C. (2014). The Cavendish Review: An Independent Review into Healthcare Assistants and Support Workers in the NHS and social care settings.

3. Methodology

Evaluation Criteria

3.1. As part of the literature review[8] of the evidence base for Values Based Recruitment (VBR), evaluation criteria with which to judge the effectiveness and efficiency of various selection methods were identified. This review of the selection methods used by Trusts was guided by this set of evaluation criteria. Appendix i details the evaluation criteria within four broad domains which are summarised below:

- **Accuracy and effectiveness**
  Evidence of reliability; evidence of validity; arrangements for on-going validation, evaluation and development are in place; susceptibility to coaching; fairness, promotes diversity/widening access; and legality.

- **Cost and efficiency**
  Scalability for high volume recruitment; efficiency; utility; and generality of use.

- **Practicalities and implementation**
  Practicality (ease of administration/efficiency); expertise required for analysis and interpretation of information generated by the tool; and ease of interpretation.

- **Stakeholder acceptance and feedback**
  Positive employee, trainee and student perceptions; generates appropriate feedback; and educational impact/value.

Scope and sample

3.2. In December 2013, NHS Employers circulated a short questionnaire to establish an overview of the work being conducted in relation to VBR at various NHS Trusts[9]. The completed questionnaires were reviewed by NHS Employers, and nine Trusts were selected to take part in follow up telephone consultation interviews between March and May 2014 to understand their approach to VBR in greater depth. These Trusts were selected in order to provide a good spread of different evidence based VBR methods, a broad coverage across professional roles assessed by the identified methods, and a diverse geographical coverage.

3.3. Of these nine trusts, eight are organisations which provide mental health services, acute hospital based care and community services and, one had explored the development of a shared service tool.

---

[8] http://hee.nhs.uk/work-programmes/values-based-recruitment/vbr-evidence/
[9] Note: for the purposes of this review only NHS Trusts were considered, and not other NHS providers such as social enterprises and private providers

3.4. All of the selection methods identified by the sampled Trusts were included in the review[10]. Table 1 below provides a summary of the identified selection methods and the types of role being recruited for in each case.

Table 1 Summary of identified selection methods and the roles recruited for

Values based screening Situational Judgement Test (SJT) N = 4	Values based structured interview N = 5	Values based selection or assessment centre N = 3
All applicants to the organisation, N = 2	All applicants to the organisation, N = 3	Senior medical professionals, N = 1
Healthcare assistants, N = 1	Allied health professionals, N = 1	Healthcare assistants, N = 1
Applicants to the learning disability service area, N = 1	Nurses, N = 1	Broad set of roles across levels, N = 1

3.5. The review accessed information from two main sources:
- Telephone consultation interviews with representatives from identified Trusts; and
- Documentation in relation to the selection methods being utilised.

Process

3.6. Telephone consultation interviews (N=9) took place with representatives from the identified Trusts which lasted between 45 – 60 minutes. A summary of the question areas used to guide the interview can be found in Appendix ii. Semi-structured in nature, the interviewees were asked to provide an overview of their approach to VBR. This was followed with broad questions around the evaluation criteria for selection that were grouped into the four categories identified from the literature review[11] of the evidence-base for VBR:

- **Accuracy and effectiveness**; for example, '*Can you provide a comment on the design or conception of the tool/method*?'
- **Cost and efficiency**; for example, '*What added value does the tool/method bring to the selection process as a whole*?'
- **Practicalities and implementation**; for example, '*What is the output of the tool/method and how is it incorporated into the broader selection process*?'

---

[10] Some Trusts are utilising more than one VBR method across their recruitment programmes.
[11] http://hee.nhs.uk/work-programmes/values-based-recruitment/vbr-evidence/

- **Stakeholder acceptance and feedback**; for example, '*What feedback have you received from all those who use the tool/method*?'

3.7. As the purpose of the interview was to gather as much detail as possible, interviewees were also encouraged to speak more openly in order for the interviewer to understand aspects of the VBR approach that were not captured by the semi-structured question areas.

3.8. In order to protect confidentiality, interviewees were informed prior to the start of the telephone interview that their input would be anonymous. Interviewees were also assured that any information they shared was to be presented back in an aggregated form and all information collected across the nine interviews would be pooled together.

3.9. To support the contents discussed during the interviews, interviewees were asked to provide any relevant material in relation to their VBR approach. Trusts differed in the material they shared, however this information generally included: Trust-specific competency frameworks, interview item-banks, scoring guides, and access to assessment platforms.

Approach

3.10. All information collected was reviewed and examined against the four categories of evaluation criteria to establish which criteria were met and the manner in which this was achieved. A full list of the evaluation criteria with descriptions can be found in Appendix i. These were cross-examined to ascertain the similarities and differences across the Trusts.

3.11. The results of this examination are presented in the next section along with the key themes which emerged from the sampled VBR methods. The manner in which aspects of the evaluation criteria were met based on the design of the sampled selection methods, and the way in which they were used at the relevant Trusts is also described.

3.12. The results of the review are presented in three subsections by selection method. Each subsection is further broken down by the four domains of the evaluation criteria as described in 3.6.

3.13. When reviewing the outcome of the evaluation criteria below, it is important to note that each criterion is not mutually exclusive. Whilst many may complement each other, it is also possible for some to be at odds with another. Thus the same piece of evidence may be used to judge the extent to which two or more evaluation criteria are met.

3.14. Complementing the results section, three case examples are given in Appendix iii, iv and v to showcase the way in which the sampled methods were used to support VBR. Each anonymised case example presents an aggregated perspective that reflects the lessons learned and strengths to take forward. The case examples also highlight the measures that were taken to address the common challenges faced by the organisations.

4. Results

4.1. The results presented below are based on the information provided from consultations with the nine identified Trusts and are therefore representative of their experiences in relation to VBR. It should be noted that due to the small sample these may not be reflective of VBR practices across **all** NHS Trusts.

Structured interviews

Structured interviews – overview of findings from the sampled NHS Trusts

4.2. The structured interview was described by Trusts as an activity which involved a meeting or telephone conversation, whereby the recruiter asks the applicant a series of questions around a predetermined set of values or competencies. Trusts differed on the degree of structure that was applied to the way in which the interviews were conducted.

4.3. The pack of materials designed to guide the interview often included a bank of different interview questions and a document designed to guide the identification of the most appropriate questions, to assess the values deemed essential in the performance of the target role. A standard scoring template, that provides a framework for the interviewer, was often used to evaluate the candidate's response and a set of scoring criteria, that determined what overall score was to be awarded based on the candidate's performance, was also common.

4.4. Trusts that used a less structured approach encouraged the interviewer to question around the candidate's values alongside their typical approach to conducting an interview. In these cases, predetermined scoring systems were less likely to be used.

4.5. The structured interview was commonly used following screening. The sampled Trusts used this method to explore how and why candidates have done certain things that demonstrated a value that is important to the organisation or role. Regardless of the degree of structure applied to the interview, candidates were typically asked to provide examples of past behaviour within real life situations. Within the structured interview, questions around values were sometimes used alongside questions assessing competencies or personal qualities.

Structured interviews – accuracy and effectiveness

4.6. As a result of unique organisational priorities, Trusts had different objectives when it came to ensuring that the structured interview tool was accurate and in identifying the form of evaluation to be carried out.

4.7. In order to demonstrate efficiency, one Trust looked at the predictive validity[12] of their structured interview. Partnering with an external provider, an evaluation study was carried out to determine the extent to which the interview scores would predict subsequent role performance. In this study, candidates who were successful in the selection process were assessed against the same set of values six months into their role. The objective of the study was to correlate candidates' scores on a set of values before and after they were recruited. Results revealed a significant positive correlation (r=.46) between the two sets of assessment scores, which indicated that the scores of that structured interview tool are predictive of subsequent role performance.

4.8. Trusts' emphasis on content validity[13] is shown through the way in which they have generally made use of subject matter experts (SMEs)[14] when developing the content of their tool. Although there is no evidence of formal evaluations of content validity, it can be inferred that the structured interviews developed in this manner are likely to be content valid as the designers had made an attempt to cover the values that the tool is intended to measure. Moreover, SMEs were involved again in generating and reviewing the context upon which the question bank is based.

4.9. Trusts' evaluation of possible **coaching effects**[15] was typically dependent on anecdotal evidence. One Trust gave the example whereby the candidate felt that the opportunity to review the question bank in advance did not improve their performance during the interview. It is suggested that the complexity of the probing questions and not knowing which areas the interviewer will probe next made it difficult for the candidate to rehearse a response. Nonetheless it is important to note that these results are based on perceptions rather than evidence; therefore in order to understand any actual effects of coaching it would be necessary to gather empirical data from appropriately designed evaluation studies.

4.10. Trusts had commonly considered **fairness** in relation to standardised administration and consistent use of the scoring scale amongst interviewers; this is discussed in greater detail along with practicalities and implementation in sections 4.13 – 4.15. Trusts generally felt that fairness could be protected by avoiding or rewording sensitive questions. In terms of monitoring fairness, there appeared to be a reliance on monitoring year-on-year recruitment outcomes at the organisational level. Using this method, Trusts generally had not found issues around diversity and inclusion; however one possible next step may be to conduct a data-driven analysis of group differences to inform possible adverse impact to subgroups. An evidence based approach can complement current measures taken to monitor recruitment outcomes at the organisational level.

---

[12] Predictive validity measures the extent to which scores of the selection method are accurate predictors of overall performance in the selection process or subsequent performance in role.
13 Content validity measures the extent to which the content of the selection method is deemed to be directly relevant to the target role by subject matter experts.
14 Subject matter experts refer to experts in a particular area of study. In the current context, they may be medical professionals to generate realistic clinical situations for behavioural assessment questions to be based upon.
15 The extent to which access to coaching taken to improve a candidate's test-taking skills may provide an advantage to a candidate's standing in the selection process.

Structured interviews – cost and efficiency

4.11. Recruiting managers generally considered structured interviews useful to help them identify the right candidate for the role and organisation. However, the use of a values based interview requires training which means that a good deal of time was spent on pre-recruitment activities for the employers to understand how the tool may be used. As a result some employers did not necessarily see structured interviews as cost efficient[16] when cost is seen as time that could be spent carrying out other aspects of their job.

4.12. A tool's **generality of use**[17] was based on judgements by the individuals responsible for implementing VBR within the organisation or Trust. Sampled organisations which used structured interviews generally asserted that they could be generalised for use in similar roles when the same set of values are assessed across the organisation. Where specific role requirements differed, relevant questions were adapted to suit the relevant context.

Structured interviews – practicalities and implementation

4.13. The structured interview was typically designed to be conducted by recruiting managers within the organisation. Whilst highly qualified in their professional area and effective in their assessment of technical competence, these managers typically lacked training to use structured interviewing techniques. Part of the shift to embed values in recruitment is to ensure that interviewers move away from traditional unstructured interviewing and learn to probe, extract and interpret behavioural evidence from a structured or semi-structured set of interview questions.

4.14. Trusts differed in the way interviewer training is carried out. Some have created a programme whereby one person from each department volunteers or is nominated to attend training. The trained person becomes the designated interviewer who will conduct all structured interviews for the department. Other Trusts have taken a train-the-trainer approach where, across a fixed period of time, all employees with management responsibilities are expected to become familiar with a values based approach to interviewing.

4.15. In addition to spending time attending interviewer training, the scope of conducting the interviews also imposed a time demand on the interviewer. Rather than forming a yes-no decision based on 'gut-feel', interviewers were required to assess the behavioural evidence in a systematic manner and to score against a standard scoring scale when using the structured or semi-structured approach. Depending on the requirement of the Trust, interviewers were also required to submit a set of interview notes to the recruitment team for administrative purposes. These activities may be seen as very time consuming.

---

[16] Cost efficiency refers to the costs involved and the time taken in developing and implementing the selection tool(s).

[17] Generality of use refers to the degree in which a selection tool used in one recruitment context can be transferred / adapted for use in another context or role.

Structured interviews – stakeholder acceptance and feedback

4.16. Given the level of involvement required, stakeholders differed in the way they reacted to the introduction of this approach. Stakeholders have generally given positive feedback if they feel that the effectiveness of the tool resulted in improvements to the recruitment process or quality of those recruited. They were also typically in favour of the wider change initiative to embed values in the organisation.

4.17. Conversely, some Trusts reported difficulty in maintaining the use of the structured interview. They report that recruiting managers do not consistently have the availability to conduct a structured interview. As a result they may become disengaged, regressing to adopting traditional interviewing approaches, or returning the responsibility to conduct the interview back to the recruitment team.

Structured interviews – key messages from the sampled NHS Trusts

- There was evidence to support that a structured interview, when designed and implemented according to best practice, can predict subsequent role performance. However Trusts differed significantly in the degree of structure that is applied to the way in which the interviews are conducted.
- Implementing structured interviews commonly involved some form of assessor/interviewer training which can be costly.
- Acceptance of structured interviews was mixed due to the complexity around their implementation; organisational engagement was a commonly identified challenge to effectively implement and sustain the use of the structured interview.
- Structured interviews were generally seen as an effective way of improving recruiting decisions; however they are also resource intensive, both in terms of the resource required to deliver the interviews and the time required to train interviewers.

Situational Judgement Tests

Situational Judgement Tests – overview of findings from the sampled NHS Trusts

4.18. Results of the review showed that when Situational Judgement Tests (SJTs) were used by Trusts, they tended to be used as a screening measure at an early stage of the recruitment process. This was to try and maximise the efficiency of recruitment processes by screening out unsuitable candidates whilst progressing a proportion of candidates that are most likely to be successful at a later stage of the recruitment process. The evidence base supporting the use of SJTs for selection was also an important factor in informing the Trusts' decision to adopt this method.

4.19. The SJTs within the sample were commonly conducted online, in an unproctored setting and at the candidate's convenience. Whilst the exact test content and wording

of the questions varied depending on the test, generally these consisted of a set of scenarios describing realistic situations faced by employees within the organisation and/or in the role. Candidates were then asked to consider a list of possible responses and to select which they were most or least likely to follow.

4.20. Candidates' responses were assessed against a predetermined key to generate a test score. This score could then be compared to a norm group to assess the candidate's relative performance against others. Finally a decision was made in relation to whether the candidate progressed to the next stage of the selection process.

4.21. When implementing SJTs, Trusts commonly drew upon the expertise of external providers who advised on the test specification, development process, implementation, and on-going evaluation. It would appear that the extent to which various aspects of best practice were met by an SJT was driven by the funding that was available to procure external expert support. Moreover, test development was a lengthy and iterative process and in practice, the contracting period may also have limited the opportunity for on-going evaluation. Taken together, it would appear that the sampled NHS Trusts have had many aspects to balance in developing their SJTs.

Situational Judgement Test – accuracy and effectiveness

4.22. On-going evaluation measures in place commonly rely on anecdotal evidence to ascertain users' perceived effectiveness. This may shed light on face validity[18], stakeholder validity[19], and content validity.

4.23. One Trust commented on an SJT's **content validity**. It was suggested that the scenarios used in that particular SJT, which was intended to be used to recruit multiple professional groups, were not reflective of the selection criteria for a particular profession. The Trust also examined the extent to which their SJT is **fair**; and results revealed significant differences across demographic groups. Taken together, these results are in line with research findings which have suggested that SJTs which are developed to be used across a broad range of professional groups are less likely to achieve good levels of validity or fairness[20].

4.24. The complex and lengthy development process for SJTs may have heightened the challenge to evaluate accuracy and effectiveness of this tool. Best practice selection process advocates the importance of conducting pilot studies before implementing a selection method for live use. The results of this pilot evaluation can then inform the possible changes that can be made to ensure on-going accurate and effective outcomes. One Trust reported that the limited availability of pilot data affected the way

---

[18] Face validity may be achieved when the selection tool content appears to be relevant towards the target role.
[19] Stakeholder validity is the extent to which stakeholders consider the content of the tool appropriate for the assessment of a particular role.
[20] 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.

that the tool was used, which in turn had an impact on the selection decisions that were made and consequently its evaluation outcomes.

Situational Judgement Test – cost and efficiency

4.25. One sampled Trust provided mixed views towards the **cost efficiency** of the SJT. It was suggested that the maximum efficiency has yet to be attained because there is uncertainty surrounding the level of confidence at which unsuitable candidates are effectively screened-out by the SJT. Across the sampled Trusts, 85 to 95% of those who complete the SJTs progressed to the next stage of the selection process. Trusts commonly reported a lack of expertise available to help determine an evidence-based cut point and as a result, a low cut point was used in order to minimise the chance of losing suitable candidates or rejecting candidates unfairly. This created a ripple effect whereby less suitable candidates were progressed, placing additional burden on subsequent selection methods to differentiate between a high volume of candidates, and on the individuals who were responsible for implementing these methods.

4.26. Users also held different views regarding the SJTs' **generality of use**. The degree of generality for SJTs depends on the specificity of the scenarios in the test. One SJT was said to be applicable to the recruitment of all roles in the organisation because the scenarios reflected work situations which may be encountered by anyone working within the organisation. In contrast, the scenarios designed to be role-specific in a different organisation were likely to be less generalisable to other roles. For example, the scenarios designed to assess empathy in healthcare workers were not necessarily considered suitable to recruit office administration staff.

Situational Judgement Test – practicalities and implementation

4.27. SJTs were generally chosen as a screening measure in order to handle high volumes of applicants and to help minimise time spent manually screening applications. This is because SJTs are designed to be automated or to require minimal administrative effort. Most of the SJTs used were hosted by an online platform and candidates were provided with an access link and a set of personal login details to complete the test in an unproctored environment during their own time. Upon completion, a computer-generated output was provided to the test administrator, who was usually a member of the programme team. The test output was typically in the form of a score which may be a percentile score to enable the comparison of the candidate's performance against a specific cohort, or an absolute score to indicate the number of questions that the candidate had answered correctly. This information was considered when deciding if the candidate would be taken forward to the next stage of the selection process.

4.28. In one Trust, a comparison between the candidate's result and the SJT cut-score was computed immediately following test completion. If the candidate's result was above the cut-score, the candidate would be directed to the next part of the online selection process.

Situational Judgement Test – stakeholder acceptance and feedback

4.29. The SJTs used by the sampled Trusts were responsible for screening out on average approximately 5 to 15% of initial applicants. As a result of the low cut point, the selection process is susceptible to false positives where candidates who are unlikely to be successful are progressed and possibly offered roles in the NHS. Stakeholder feedback supported this notion as recruiting managers asserted that the quality of candidates assessed had not improved following the changes. More importantly, the recruiter's load had not actually lessened when it came to implementing subsequent selection measures on a larger scale.

Situational Judgement Tests – key messages from the sampled NHS Trusts

- Situational Judgement Tests (SJTs) are often chosen by Trusts as a way of improving the efficiency of the selection process
- The SJTs used by the sampled Trusts were responsible for screening out on average approximately 5 to 15% of initial applicants. Maximum efficiency was yet to be attained because there was uncertainty surrounding the level of confidence at which unsuitable candidates are effectively screened-out by the SJT. Trusts commonly reported a lack of expertise available to help determine an evidence-based cut point and as a result, a low cut point was used in order to minimise the chance of losing suitable candidates or rejecting candidates unfairly.
- Developing and implementing an SJT according to best practice is time consuming and can be resource intensive, representing a significant investment for an individual Trust. If efforts and resources were combined, it is more likely that valid and fit-for-purpose SJTs could be developed.
- As automatically scored tools, SJTs can reduce the need for resources to be spent manually screening applications and as such are considered to offer a useful screening mechanism especially for roles with high volumes of applicants. However, the results from this review highlight that in order for Trusts to be confident in using SJTs to screen out significant numbers of applicants, there is a necessity to undertake robust pilot evaluation prior to live use, and to put in place measures to enable on-going evaluation. In the absence of such evaluation, SJTs are less likely to have the desired impact.

Selection centres

Selection centres – overview of findings of the sampled NHS Trusts

4.30. Results of the review demonstrated that selection centres (SC) used by the sample Trusts typically comprised a series of selection exercises to assess a range of attributes including a candidate's values. The selection exercises often included role-play exercises, presentations, written tasks or group activities. They were designed to assess different values, competencies or personal qualities which were assessed by multiple assessors. The SCs were either conducted as half or full day events depending on the objective of the recruitment process and the resources available.

4.31. In addition to generating scores based on candidates' performance at each selection exercise, typically a broader scoring algorithm was applied to tabulate candidates' performance across the SC to infer overall suitability for the role.

4.32. It should be noted that whilst three Trusts reported the use of SCs as part of their VBR approach, minimal information were drawn from two examples. For this reason, the results below focus predominately on one Trust.

Selection centres – accuracy and effectiveness

4.33. Targeting widespread issues around poor quality of recruitment and low retention rates, an SC was designed at one Trust to enable the assessment of values that are specific to a senior medical consultant role.

4.34. To enable any changes following the introduction of the SC to be assessed, the Trust took a baseline measure of workforce metrics[21] prior to the implementation of the SC. Following its implementation, the same workforce metrics were measured again to enable a comparison. The correlational study found that the SC resulted in improved recruitment outcomes for the target role. These improvements included higher retention rates, reduced turnover rates and lower spending on recruitment overall. Such improvements are assumed to be antecedents for achieving improvements in the quality of services delivered including enhanced patient outcomes.

Selection centres – cost and efficiency

4.35. There was a large cost to implement the SC due to the level of resources that were required from the programme team, administrators and recruiting managers, as well as venue hire and logistical coordination.

4.36. As a result of the time and financial resources required, SCs are often used only for the recruitment of senior level roles. This therefore reduces the likelihood of generality of use, even if it may be possible to adapt the same SC design for use in the selection of a different role.

Selection centres – practicalities and implementation

4.37. Due to the complexity of SC exercises and overall design, the programme team or the designers of an SC typically assumed a high degree of involvement in the running of the event. Recruiting managers were also often involved as assessors to give a balanced perspective on the candidates' performance. The assessment at an SC relies on the assessor's judgement to evaluate the candidate's performance based on a set of standard scoring criteria.

---

[21] Workforce metrics refer to indicators which may suggest that a target recruitment strategy is effective. Retention rates, turnover rates, cost per recruit, return on investment are examples of workforce metrics.

4.38. Assessors were generally required to attend assessor training prior to the event where they were briefed on the exercises, the behavioural evidence to be extracted and how performance was to be scored. Following training, they were typically required to attend the full duration of an SC, which may include a calibration session where assessors deliberate their assessments. Calibration sessions were generally led by the programme team which played a role in coordinating scoring decisions. It can be seen that the preparation and running of an SC places a demand on time for a large number of people.

Selection centres – stakeholder acceptance and feedback

4.39. Given that an SC is resource intensive to implement, one might expect recruiting managers to become disengaged – reactions that are similar to the structured interview. However it was found that recruiting managers and administrators have typically responded favourably and report finding that it is an effective manner to recruit for values. The differences in reaction may be the result of greater acceptance for the wider initiative to embed values that is unique to the organisational context.

Selection Centres – key messages from the sampled NHS Trusts

- There was evidence to demonstrate improvements to recruitment outcomes following the implementation of a Selection Centre (SC). Evaluations of reliability and validity could provide further support for the accuracy and effectiveness of this method
- SCs were generally considered resource-intensive to implement and the time and financial costs required to design and implement SCs were commonly cited reasons for using this approach for the selection of senior level roles only.
- Despite the labour intensive nature of SCs, stakeholders had generally responded favourably to their introduction as they were seen as adding significant value to a selection process.

---

5. Summary of Key Learning Points

5.1. This review has provided an overview of selection methods currently used by nine Trusts to recruit for values. This information may be used to support the development of guidance to help Trusts implement VBR in the future or enhance existing VBR practices. Below is an outline of the key learning points applicable to the wider Trust community, based on the information provided by the sampled Trusts. This is structured first by selection method (structured interviews, Situational Judgement Tests and selection centres) and then by the evaluation criteria (accuracy and effectiveness; cost and efficiency; practicalities and implementation; and, stakeholder acceptance and feedback).

Summary of key learning points – by selection method

Structured interviews

5.2. Organisational engagement and stakeholder consultation are important aspects when developing a structured interview. The involvement of stakeholders early on in the development process helps Trusts to ensure that the content of the interview question-bank is relevant to the target role, profession or level as well as helping to build stakeholder engagement in the process.

5.3. The effectiveness of the structured interview is reliant on the manner in which the interviews are conducted and scored. As a result, emphasis should be given to the interviewer training sessions to enable interviewers to apply consistent and standardised interviewing, probing and scoring techniques.

5.4. The requirement for fit-for-purpose interviewer training should be balanced with a consideration of the time demand that is placed on the recruiting manager.

5.5. Feedback from candidates should be gathered to provide insight into their perceptions of the interview methodology and the appropriateness of the questions. Furthermore, where possible, information regarding the candidate's demographic details should be gathered to enable the analysis of sub-group differences. This could be gathered during the pilot phase, and from current role-holders, in order to avoid sensitivity but without compromising its usefulness. When gathered, whether during the piloting or live assessment, it should be made clear to candidates that this is being used for monitoring purposes only and will not inform the selection decision. Such questions should also be voluntary.

5.6. Local measures should be in place to enable on-going evaluation of this method. In addition to receiving stakeholder feedback to infer the level of acceptance and perceived effectiveness, interview data may undergo analyses to determine the extent to which the method is accurate and effective. This would provide robust evidence for the Trust to support its continued use and identify possible areas of refinement.

Situational Judgement Tests

5.7. The development of Situational Judgement Tests (SJTs) can be time consuming for Trusts and require expert input to ensure that they are appropriate and fit-for-purpose. Generally, SJTs which are developed through close consultation with SMEs, from within the Trusts, are likely to be better received.

5.8. Ideally, SJTs should be subject to piloting prior to live use in order to enable psychometric evaluation of the tool's accuracy and effectiveness, focusing on measures of validity and fairness as a priority. Analysis may also include data modelling to determine a fit-for-purpose scoring algorithm and appropriate cut scores. This might provide the evidence required for a Trust to support the continued use of the tool as a mechanism for sifting out a greater number of applications.

5.9. To ensure that the content of SJTs remains up to date and the security of items are not compromised through continued use, a process to continuously develop new content for the tool should ideally be in place.

5.10. SJTs are often administered online. This may present challenges in terms of integrating the test into existing Trust application systems. If the SJT is to be administered online, appropriate checks should also be put in place to deter cheating. This might include asking applicants to complete the SJT in a proctored setting or implementing a follow up verification test later in the selection process.

5.11. Feedback from applicants should be gathered to provide insight into their perceptions of the SJT methodology and the appropriateness of the scenarios. Furthermore, where possible, information regarding the applicant's demographic details should be gathered to inform the analysis of sub-group differences. When gathered, it should be made clear to applicants that this is being used for Trust monitoring purposes only and will not inform the selection decision.

5.12. This review highlights some of the challenges associated with individual Trusts attempting to undertake the development of SJTs in isolation. Such challenges are largely attributed to a lack of resources available locally to invest in development, refinement, piloting and evaluation in accordance with best practice. An approach whereby a number of Trusts work as a consortium to pool resources and invest in on-going development and evaluation of SJTs may be more likely to result in the development of robust and fit-for-purpose tools.

Selection centres

5.13. The development of a selection centre (SC) should ideally be preceded by a job/role analysis to identify and customise the values that are required for success in the role based on the NHS Constitution. It should also inform a set of scoring criteria to guide the evaluation of performance within each individual exercise and the overall SC.

5.14. Stakeholder consultation should provide a role-specific context upon which individual exercises and the overall SC could be based.

5.15. The overall effectiveness of an SC, in any Trust, is reliant on the extent to which assessors are able to extract and score behavioural evidence. It is therefore important that assessors are given adequate training prior to the 'live event'.

5.16. The requirement for adequate assessor training should be balanced with a consideration for the time demand imposed on the recruiting manager.

5.17. The benefits of an SC design may need to be balanced with a consideration of the cost involved to design, implement and evaluate the approach. There is typically greater justification for the recruitment of senior roles using this approach.

Summary of key learning points – by evaluation criteria

Accuracy and effectiveness

5.18. The evaluation of effectiveness is driven predominately by stakeholder perceptions. There appears to be limited opportunity for the Trusts to conduct full (psychometric) analysis to evaluate a selection method's reliability, validity, fairness, coaching effects, and legality. Common challenges to undertake these analyses were identified as the lack of expertise, financial resources and the availability of data.

5.19. The evaluation of group differences requires the collection and analysis of demographic data, such as gender, age or ethnicity, in relation to performance on the assessment. As these may be considered as sensitive personal information, organisations may opt not to request them in order to observe, or to maintain an image of, equality and diversity protection. Thus Trusts may be faced with a conflict when the collation of demographic information is a necessary step to evaluate and refine a selection method in relation to fairness.

5.20. Trusts' common indicators of a selection method's immediate success were taken as the reduction in time spent on recruitment activities and the costs saved as a result. Evaluations of effectiveness generally looked at the extent to which time and / or financial efficiency were achieved. Based on the results of the evaluation, some decisions were made with regards to whether funding for the local VBR approach may be continued.

5.21. Trusts were also interested in the long-term effectiveness of a selection method(s). As a consequence, some Trusts are conducting studies in order to uncover evidence of validity and ensure that a measure of on-going evaluation is in place.

Cost and efficiency

5.22. Trusts' evaluation of costs and efficiency commonly focuses on the extent to which the (perceived) time taken to recruit and develop the right employees is reduced. The selection methods implemented are likely to be seen as successful if they are perceived to be able to alleviate the time demand that is placed on the recruiting

managers. It is expected that if role performance improves, the demand on recruiting managers to retrain and rehire may be reduced.

5.23. There has been limited attention given to establishing workforce metrics to conduct statistical analysis on scalability, efficiency, and utility. The results presented are drawn primarily from anecdotal evidence where Trusts provided feedback on perceived outcomes when a values based selection method was used.

5.24. The evidence suggests that employers hold mixed views toward whether the implemented method has improved efficiency and reduced their workload.

Practicalities and implementation

5.25. The way in which selection tools are implemented is an important consideration for the Trusts. As implementing a values based approach to recruitment is often associated with broader organisational change initiatives, the inclusion of stakeholders during design and implementation is particularly valuable.

Stakeholder acceptance and feedback

5.26. Stakeholder feedback is commonly drawn from staff involved in the delivery of the selection tools. Using methods such as focus groups, Board debriefing meetings or informal conversations, Trusts provide feedback on various aspects such as content relevance, ease of use, or tool effectiveness. Understanding candidates' perceptions of the selection methods is a useful way of monitoring the face validity and perceived fairness of the tool as well as identifying possible improvements which could be made to the way it is administered. However, the majority of the Trusts did not report collecting feedback from candidates.

5.27. The way in which Trusts provide assessment feedback to candidates varies but typically falls under two categories. First, feedback may be given to successful candidates; this is generally given from a development perspective and can be linked to subsequent appraisal measures. Second, automated feedback may be given to candidates to inform whether their performance on a screening measure had satisfied the cut-score to enable progression to the following stage of selection.

6. Recommendations

6.1   The information presented in this report provides insight into the way in which VBR is currently being practised within Trusts. The following recommendations are suggested in response to the evidence which has emerged through this review and alongside the evidence from the previous literature review[22].

Recommendations for the VBR Programme

- **Recommendation 1a: Generate practical case study material, guidance documents, resources, training and support to be shared with employers to coincide with the launch of the VBR framework in October 2014.** Whilst some resources may be generic and accessed by all Trusts, some materials may benefit from being tailored for specific Trusts to account for the differences in how recruitment is delivered and their specific local challenges and needs. Resources should also be derived using the available evidence base and take into account the common approaches currently used. Resources should be available to coincide with the launch of the VBR framework in October 2014.

- **Recommendation 1b: Design mechanisms to further enhance understanding of NHS Constitution values within Trusts to help ensure their more consistent promotion and application, and provide resources, such as the NHS Employers Values Mapping tool, to enable organisations to map their own local values with those of the NHS Constitution.** Provide resources to enable organisations to map their own local values with those of the NHS Constitution, for example through the NHS Employers Values Mapping tool.[23]

- **Recommendation 1c: Coordinate resources for Trusts to access to support the development, implementation and evaluation of VBR.** This will help to ensure that those responsible for implementing VBR within Trusts have access to appropriate guidance and that the evidence-base for VBR remains up to date in the light of emerging findings from both academia and practice.

Recommendations for individual NHS Trusts when undertaking VBR

- **Recommendation 2a: Conduct a job/role analysis to systematically identify the values to be assessed in relation to the NHS Constitution.** The basis of an effective selection method is a thorough analysis of the relevant values, competencies and skills that are required for successful performance in the target role. As suggested in recommendation 1b, Trusts may map their local values against those of the NHS Constitution. Based on a thorough understanding of the key requirements for success, the most appropriate selection method to assess for the identified values may be identified. The outcomes of the analysis may also

---

[22] http://hee.nhs.uk/work-programmes/values-based-recruitment/vbr-evidence/
[23] NHS Employer's Values Mapping Tool can be accessed via http://www.nhsemployers.org/case-studies-and-resources/2014/04/values-mapping-tool.

inform the scoring criteria and weight that is given to the selection method within the wider recruitment process.

- **Recommendation 2b: Incorporate stakeholder consultation when developing the selection method to maximise organisational engagement and to ensure a high degree of relevance to the role.** Selection methods should ideally be developed through close consultation with Subject Matter Experts (SMEs) from within the recruiting organisation. Selection methods developed in this way are more likely to be well received by recruiting managers and candidates.

- **Recommendation 2c: Provide training to assessors to ensure appropriate use of the selection method and ensure that evidence of a candidate's values may be extracted and scored in a standardised way.** Some selection methods, such as the structured interview, may require the assessor to formulate a judgement on the candidate's performance. It is important that assessors receive prior training to ensure that behavioural evidence may be extracted and scored in a standardised manner to ensure fairness, reliability and validity. The design of the training process should be balanced with the consideration of the demand imposed on the recruiting manager.

- **Recommendation 2d: Implement on-going evaluation measures to ensure continued improvements in selection method(s) including monitoring of possible sub-group differences.** Psychometric evaluation studies can provide greater rigour when monitoring the effectiveness and efficiency of the selection methods. Evaluation can also be undertaken to examine possible sub-group differences in performance at selection (for example on the basis of age, race or gender) to contribute to ensuring the fairness of a selection system. This can inform continual improvement and help to refine the method(s). Nonetheless, lack of expertise and/or financial resources were commonly identified as challenges to conducting evaluation. In this way, Trusts are likely to benefit from support and/or training to carry out evaluation to support the emerging evidence-base for VBR.

Appendices

Appendix	Title	Page
i	Evaluation criteria for selection methods	28
ii	Semi-structured interview for Trusts – question areas	31
iii	Case example 1 – structured interviews	32
iv	Case example 2 – Situational Judgement Tests	36
v	Case example 3 – selection centres	41

Appendix i: Evaluation criteria for selection methods

Category	Evaluation Criteria	Description	How can this be evidenced?
Accuracy and effectiveness	1. **Evidence of reliability**	*A selection method is reliable if it is consistent in how it assesses candidates under varying conditions. For example, it should not make a difference if a candidate sits the test in the morning or afternoon.*	• *Psychometric evaluation by experts*
	2. **Evidence of validity**	*The selection tool measures what it claims to measure, it should be relevant, precise and accurate.*	• *Psychometric evaluation by experts*
	3. **Arrangements for on-going validation, evaluation and development are in place**	*Best practice selection is an iterative process, starting with a job analysis to define the selection criteria. After selection has taken place, the predictive validity of various selection tools can be evaluated. Results from validity studies are then used to review the original selection criteria and choice/design of selection methods. Information here can be used to make continual improvements and help to develop the process to optimise selection decisions.*	• *Appropriate data is collected* • *Validation data is analysed by experts in selection*
	4. **Susceptibility to coaching**	*The extent to which access to coaching taken to improve a candidate's test-taking skills and provide an advantage to a candidate's standing in the selection process.*	• *Comparison group research studies*
	5. **Fairness, promotes diversity/ widening access**	*This is based on three principles; 1) valid selection criteria;* *2) accurate and standardised administration by trained staff;* *3) monitored outcomes and meets equalities impact assessments.*	• *Evaluation questions posed to candidates.* • *Analysis of sub-group differences*
	6. **Legality**	*The extent to which the design of a selection process and the decisions generated is legally defensible. Selection processes that are perceived as unfair are more likely to result in legal case initiation.*	• *HR experts in employment law*
Cost and efficiency	7. **Scalability for high volume recruitment**	*The extent to which a selection process can be scaled up or down and remains efficient and effective for different volumes of applicants.*	• *Data modelling with interpretation of costs of implementation and validity of selection methods*

	8. **Efficiency**	*The costs involved and the time taken in developing and implementing the selection tool(s).*	• *Analysis of costs by recruiters and managers*
Practicalities and implementation	9. **Utility**	*The costs involved and the time taken to develop more accurate adequate procedures need to be balanced with the potential benefits (e.g. improved performance)*	• *Statistical analysis of the predictive validity a selection tool adds to the accuracy of selection decision-making, compared to the costs to design and implement the tool (using established utility calculation methods)*
	10. **Generality of use**	*The degree to which a selection tool used in one context can be transferred/tailored for use in another context or role*	• *Judgement by recruiters*
	11. **Practicality (ease of administration/ efficiency)**	*The procedures should be acceptable within the organisation and capable of being implemented effectively. Those responsible for administering the procedures may need to be trained.*	• *Judgement by recruiters and administrators*
	12. **Expertise required for analysis & interpretation of information generated by the tool**	*Some selection tools (for example personality tests) require an appropriately trained individual to administer, score and provide feedback. Similarly, assessors in selection centres must also be appropriately trained in how to evaluate a candidate in a work sample test for example.*	• *For psychometric tools there are specific licensure guidelines (e.g. from the British Psychological Society)*
	13. **Ease of interpretation**	*The degree to which the information generated by the selection tool provides clear and appropriate information relating to a candidate's competence and aptitude for the role.*	• *Judgement by recruiters and stakeholders*
Stakeholder acceptance and	14. **Positive employee/trainee/student**	*The extent to which employees/trainees/ students react positively to the selection process and each selection method*	• *Evaluation questionnaires of candidate perceptions*

feedback	perceptions	*within that process. Positive perceptions will result in the candidate being more attracted to joining the organisation*	
	15. Generates appropriate feedback	*When using selection tools, for example personality assessments, it is good practice to ensure that candidates receive appropriate and useful feedback.*	• *Evaluation questionnaires of candidate perceptions* • *Recruiter's judgements*
	16. Educational impact/value	*The extent to which candidates obtain useful information to inform their future education, learning and development.*	• *Evaluation of candidate and employer perceptions*

Appendix ii: Semi-structured interview for Trusts – Question areas

1. **General**
   - Can you tell me about the Trust's approach to Values Based Recruitment – particularly any tools that you may be using to assess values?

   - Which recruitment process(es) is the tool currently being used within? (May group these into broad functions, for example Medical staff, Support staff, Allied Health Professionals etc.; should also link this into the scale of uses depending on the role)

   - Thinking back to the design / conception phase, what was the rationale for selecting this tool? What other tools/methods were considered? What were the reasons for discounting other approaches? What were the most important factors which influenced the decision about which approach to take?

2. **Accuracy and effectiveness**
   - Are you able to comment on the design or conception of the tool? What was this based on? What factors were taken into account to ensure a quality end product?

   - Has there been any analysis/research undertaken to explore the impact of the tool on quality of those recruited? (For example, the extent to which it predicts future behaviour?)

   - Has there been any analysis/research undertaken to explore the fairness of the tool? (For example, any differences across demographic groups? Any issues in relation to coaching?)

3. **Cost and efficiency**
   - Are you able to give a comment regarding the cost efficiency of the tool? What is the pricing structure of the tool (for example, one-off development cost, price per use?) Did the cost have an impact over its procurement?

   - What added value does the tool bring to the selection process as a whole?

   - In your view, can the tool be used for other purposes within (insert name of Trust)? (For example, recruitment for other roles, employee development, training?)

4. **Practicalities and implementation**
   - Who is responsible for administering the tool? Can you describe to me how the method is administered?

   - How do administrators / assessors prepare themselves before using the tool?

   - What is the output of the tool? How is the output used in the selection process(es)?

   - Is there a specific person who is responsible for the interpretation of results?

5. **Stakeholder acceptance and feedback**
   - What feedback have you received from all those who use the tool?

   - How is the tool perceived generally by the Trust, stakeholders, and/or candidates?

Appendix iii: Case example 1 – structured interview

1.1. Five Trusts reported the inclusion of the structured interview in their selection processes. The interviews were generally designed as a set of broad interview questions with the intention that the user, usually the recruiting manager, would identify a subset of questions relevant to the target role and adapt the wording to enable the assessment of the pre-determined values or qualities required for success in the role.

1.2. The Trusts began to adopt the structured interview as one way to address organisational issues such as high turnover, low motivation and poor job performance. Traditionally, Trusts made recruitment decisions based on information from CVs, personal references, or responses from unstructured interviews. There was growing recognition that these selection methods profile the candidate's formal qualification, technical competence and past experience. However they provide minimal information with regards to the candidate's personal values or behavioural preferences that can indicate how they may perform the role. To address this, Trusts began to explore the possibility of assessing non-technical qualities to complement existing measures of professional competence.

1.3. Unique organisational contexts gave Trusts different reasons to choose a structured interview tool, however people engagement was a commonly identified driver. Trusts believed that adopting a structured interview can create opportunities for stakeholder engagement across stages of development, implementation and evaluation. More importantly, the tool itself was considered to create structured opportunities for recruiting managers to get to know potential employees, whilst enabling candidates to demonstrate their non-technical qualities.

1.4. The development of a structured interview was commonly nested as one project within a larger organisational development (OD) initiative to introduce VBR to the organisation. For this reason, efforts to embed values across the organisation typically began in-house with variant forms of a job analysis[24] to identify the necessary qualifications, competencies, and values that are essential for effective performance across various job families. With stakeholder involvement, these competencies and values were then threaded throughout various OD interventions and talent management strategies, which included recruitment processes and selection method design.

Developing a structured interview

1.5. All programme teams had a great deal of involvement in developing their structured interview. For some, the tool was designed completely in-house while others had partnered with an external provider. Commonly led by the programme team, the development of the tool began with a stakeholder consultation event. This event

---

[24] A job analysis is a systematic analysis of the relevant knowledge, skills, abilities and attributes associated with performance in the target role. This is the first stage of designing a selection process or method in accordance to best practice.

served as one way to elicit the content of the tool from stakeholders and subject matter experts (SMEs)[25], and in doing so, opened up an opportunity to disseminate the broader OD initiative.

1.6. Trusts engaged with their stakeholders in different ways however their aims were similar whereby SMEs were given the set of organisation-specific values and asked to describe the way in which each value may be demonstrated within their department, profession, or role level. Their input formed the basis of the interview questions and scoring criteria that were designed to be broadly relevant across the various job families in the organisation. In this way the interview questions were broadly aligned with the values of the organisation, thus demonstrating content[26] and face[27] validity.

Using a structured interview

1.7. Engaging a potential employee in a conversation to elicit and enable assessment of non-professional attributes is a core objective of a structured interview. As with any other selection method, the effectiveness of a tool is partly dependent on the way it is implemented. Trusts recognised this and accordingly interviewer training sessions were in place to teach effective interviewing and probing techniques. Recruiting managers were also taught how to extract information from interview data and how scoring should be conducted. Examples of sensitive questions[28] that may impact on fairness were identified and recruiting managers were advised to avoid or be cautious around these.

1.8. Recruiting managers reacted to the training sessions differently. Whilst some welcomed the opportunity to up-skill their interviewing technique, others had difficulty coping with the additional demand on their time. Some Trusts asserted that the latter group demonstrated lower levels of engagement and commitment to the OD initiative as a whole. Within these Trusts, recruiting managers abandoned the structured interview approach altogether which had a negative effect on the wider OD initiative.

Evaluating a structured interview

1.9. As each Trust was faced with the challenge to minimise cost, there was pressure to demonstrate that their structured interviews were as effective as possible. Accordingly evaluation outcomes were important deciding factors for continued funding. Nonetheless the time and manner in which evaluation was conducted differed depending on each Trust's unique OD context.

---

[25] Subject matter experts refer to experts in a particular area of study. In the current context, they may be medical professionals to generate realistic clinical situations for behavioural assessment questions to be based upon.
[26] Content validity is the extent to which the content of the selection method is deemed by SMEs to be directly relevant to the target role.
[27] Face validity is the extent to which the selection tool content appears to be relevant to the target role.
[28] Sensitive interview questions are questions that require the candidate to disclose personal information that should not form the basis for recruitment decisions.

1.10. In terms of how evaluation was conducted, a Trust's own evaluation of the structured interview typically relied on anecdotal evidence to infer effectiveness and cost efficiency. Whilst most Trusts acknowledged a lack of expertise to conduct an evaluation, some had partnered with external providers to carry out larger scales evaluations studies examining the predictive validity[29] of the interview or as an on-going measure to justify spending.

1.11. Preliminary stakeholder feedback showed mixed results. Initial stakeholder feedback revealed positive reactions to the outcomes of the structured interview. As well as favourable feedback from recruiting managers where they felt that they are able to 'get the right people the first time', ward managers also felt that the new recruits are more prepared for the job. Many Trusts have also noticed significantly lower levels of turnover since implementing a structured interview. However on the flip side, training of the interviewers was identified as a key challenge. There were concerns around recruiting managers' availability to conduct structured interviewing and it was speculated as a reason for some of the negative reactions towards the tool.

1.12. Only one set of data-driven results was available at the point of this review, while others are in the piloting or early stages of implementation. The available data suggested that the structured interview was a valid predictor of subsequent in-role performance at that particular Trust, and as a result the tool is being piloted across a wider range of job families.

Next steps

1.13. Structured interviews had improved recruitment outcomes across the Trusts that were sampled. They have also demonstrated varying levels of effectiveness in meeting their objectives. In addition to improving the structured interview further, specifically to manage the impact that interviewer training may have on engagement, some of these Trusts were also investigating the possibility of pairing the structured interview with a screening measure to be completed ahead of the interview. This was partly to alleviate the burden placed on the interview itself and the interviewers' time to assess a relatively high volume of candidates. It is expected that a screening measure may streamline the broader selection process by focusing resources on candidates with the highest potential for success at the structured interview stage.

Key learning points and issues for consideration

1.14. Organisational engagement and stakeholder consultation are important aspects when developing structured interviews. The involvement of stakeholders early on in the development process lends support to the running of the wider OD programme whilst ensuring that the content of the interview question-bank may be highly relevant to the target role, profession or level.

---

[29] The predictive validity measures the extent to which outcomes of the selection tool are accurate predictors of subsequent performance; which may be the overall performance of the selection process or post-recruitment in-role performance.

1.15. The effectiveness of the structured interview is reliant on the manner in which the interviews are conducted and scored. As a result, emphasis should be given to the interviewer training sessions to equip interviewers with standardised interviewing, probing and scoring techniques.

1.16. The requirement for fit-for-purpose interviewer training should be balanced with a consideration of the time demand that is placed on the recruiting manager.

1.17. Feedback from candidates should be gathered to provide insight into their perceptions of the interview methodology and the appropriateness of the questions. Furthermore, where possible, information regarding the candidate's demographic details should be gathered to inform the analysis of sub-group differences. As candidates are more likely to offer personal information in low stakes situations, efforts to gather this data may be best placed during the development stages of the selection method. When gathered, it should be made clear to candidates that this is being used for monitoring purposes only and will not inform the selection decision. Such questions should also be voluntary.

1.18. Measures should also be in place to enable on-going evaluation of the structured interview. In addition to receiving stakeholder feedback to infer the level of acceptance and perceived effectiveness, interview data may undergo psychometric analyses to determine the extent to which the interview is accurate and effective. This may provide robust evidence to support the continued use of the method and identify possible areas of refinement.

Appendix iv: Case example 2, Situational Judgement Tests

1.1. Four Trusts reported the use of a Situational Judgement Test (SJT) as a screening tool used at an early stage of the selection process. Trusts utilised SJTs to assess for different roles. For some, an SJT was being used generically for the assessment of broad professional roles. For others, it was designed for the assessment of a specific role in a support function that typically attracts a high volume of applicants.

1.2. The Trusts were interested in developing a values based SJT so that selection decisions could be based on more than one source of information; more specifically, to enable the consideration of a potential employee's values. Instead of or in addition to using an interview tool, four Trusts considered the development of an SJT as a means to streamline selection processes for high volume roles. It was expected that the SJTs would be used as an automated screening tool placed at early stages of the selection process to sift through the highest ratio of applicants.

Developing a bespoke online Situational Judgement Test

1.3. The development of an SJT was commonly nested within broader organisational development (OD) projects. The development typically began with the identification of a set of organisation-specific values using variant forms of a job analysis methodology. These identified values were taken to inform the SJT specification that guided later stages in the test development process.

1.4. Due to the scale of these OD initiatives and the level of expertise required to develop a test, all the Trusts partnered with external providers who took the lead in the test development process as they provided specialist psychometric expertise.

1.5. Stakeholder engagement was a common theme in each of the separate development processes. The level of stakeholder involvement differed across the Trusts, however the consultation process was similar whereby focus group events were conducted to draw on stakeholders' input to inform the SJT assessment criteria. As there was variability in the way in which these assessment criteria were taken forward, two different outcomes are described below.

Developing a bespoke online Situational Judgement Test – designing test content from scratch

1.6. Following the focus group event to determine the assessment criteria, one Trust engaged in further consultation events to draw on a group of subject matter experts'[30] (SMEs) technical input on the content of the SJT. This provided a role-specific context for which scenarios can be designed to assess values rather than procedural knowledge. In this way the scenarios designed were bespoke and unique to the

---

[30] Subject matter experts refer to experts in a particular area of study. In the current context, they may be medical professionals to generate realistic clinical situations for behavioural assessment questions to be based upon.

Trust's organisational context which increased the likelihood of achieving stakeholder acceptance[31] and content validity[32].

Developing a bespoke online Situational Judgement Test – drawing from an established scenario bank

1.7. Taking on a different approach, another Trust appeared to have given more weight to the competency framework already established by the chosen provider. Using this approach, the provider's competency framework is assumed to be an appropriate match to the Trust's and/or the NHS Constitution values. Based on this assumption, the pre-determined values were mapped to the provider's competency framework in order to draw on the established scenario bank and identify appropriate scenarios to build the content of the SJT. The benefit of this approach is the confidence that the established competency framework is a robust and valid one, and that the scenarios drawn have been validated psychometrically. Whilst this approach minimises the resources required for test development, it may reduce the opportunity for stakeholders within the Trust to contribute to its design. In this way, the validity of the tool may be compromised by the ease of development.

Piloting and evaluating Situational Judgement Tests

1.8. As part of the test development process, best practice suggests that assessments undergo review and piloting prior to their live use[33]. Based on the results of this review stage, potential issues like adverse impact on different sub-groups may be identified and the feedback received can inform areas of possible refinement. Trusts appeared to take different approaches to piloting.

1.9. As each sampled Trust was faced with the challenge to minimise cost, there was pressure to ensure that the SJT was effective in meeting its objectives. In this way the outcome of evaluation studies were important deciding factors for continued funding for all of the Trusts. Nonetheless, the time and manner in which evaluation studies were conducted differed depending on each Trust's unique circumstances.

Piloting before live use

1.10. One Trust was keen to demonstrate effectiveness early on in the development process due to the funding structure, and as a result a pilot study is currently being carried out prior to live use. This study aims to establish the predictive validity[34] of the SJT and results of this pilot will: 1) Establish a data-driven cut score, 2) Inform a decision on

---

[31] Stakeholder acceptance is the extent to which stakeholders consider the tool appropriate for use in a particular selection process.

[32] Content validity is the extent to which the content of the selection tool is deemed by subject matter experts to be directly relevant to the target role(s).

[33] Patterson F. (2012). Selection Methods. In Cleland, Dowell, McLachlan, Nicholson, & Patterson, *Identifying best practice in the selection of medical students*. Research report to the General Medical Council.

[34] The predictive validity measures the extent to which outcomes of the selection tool are accurate predictors of subsequent performance; which may be the overall performance of the selection process or post-recruitment in-role performance.

whether the SJT will be used in the recruitment of other roles, and 3) Enable a decision on whether additional funding will be allocated to increase the item bank.

Evaluation following implementation

1.11. Taking on a different approach, one Trust had conducted its evaluation of their SJT based on the data collected from the live full scale assessment. The objective of the evaluation was to justify the costs spent.

1.12. Initial content review by SMEs revealed that the provider's competency framework was not a close match to the organisation's values. In particular, content reviewers felt that the scenarios used were not a realistic reflection of the target role (or many other volume roles within the organisation). Adding to the limited content validity, it was subsequently found in an analysis of group differences that a subgroup was disadvantaged when assessed by this particular SJT. Arguably the extent of the adverse impact may have been minimised if given the opportunity to pilot and refine prior to full implementation.

Assigning cut-scores

1.13. Pilot studies can provide the added benefit of informing a cut-score or scoring algorithm based on data modelling33. The assignment of weighting to various assessment criteria or selection exercises within the process can draw on the results of pilot scores. As the sampled Trusts had not had access to this information at the point of implementing an SJT, they have approached this with caution. Careful to not reject potentially suitable applicants on the basis of the SJT alone, approximately 85 to 95% of all applicants were put forward to the next stage of the selection process. In this way, the SJT was only minimally effective in reducing the applicant pool prior to shortlisting. Nevertheless the Trust expected that further data will provide more clarity to enable a larger proportion of unsuitable applicants to be screened out on the basis of their performance on the SJT.

Administering the Situational Judgement Test

1.14. Some Trusts reported challenges around the online administration of an SJT. Assessment users without an existing IT platform would often enjoy the convenience that assessment providers offer when it comes to launching an assessment for live use. However, for organisations with an existing applicant tracking system, this poses a challenge for the two systems to be bridged. Candidates may struggle to navigate the various systems when they are given several different sets of login details to access multiple systems to complete the various components of the application process.

1.15. Aside from challenges to integrate an SJT administration platform, Trusts also have generally reported a decrease in the number of applications made since the introduction of the SJT. Further data is required to explore the reasons for this reduction however it can be speculated that this is likely to be due to a combination of two factors. Firstly, simply introducing another process in the selection process may

discourage a sample of applicants due to the additional effort required. Secondly, the SJT serves as a 'self-selection tool' whereby through the course of considering the scenarios, applicants self-reflect on their suitability for the role and some subsequently chose not to continue with their application.

Next steps

1.16. Test design and evaluation is a lengthy process and as the sampled Trusts illustrated, it is an on-going and iterative one. All Trusts are currently working on evaluating and fine tuning their screening SJTs further.

Key learning points and issues for consideration

1.17. The development of SJTs can be time consuming and require expert input to ensure that they are appropriate and fit-for-purpose. Generally, SJTs which are developed through close consultation with SMEs from within the recruiting organisation are likely to be better received.

1.18. Ideally, SJTs should be subject to piloting prior to live use in order to enable psychometric evaluation of the tool's accuracy and effectiveness, focusing on measures of validity and fairness as a priority. Analysis may also include data modelling to determine a fit-for-purpose scoring algorithm and appropriate cut scores. The piloting, evaluation and analysis stages could be factored in to the development and on-going evaluation phases of the tool. This is likely to provide the evidence to support the continued use of the tool as a mechanism for sifting out a greater number of applications.

1.19. To ensure the content of SJTs remains up to date and the security of items are not compromised through continued use, an on-going process to refine the tool should ideally be in place.

1.20. SJTs are often administered online. This can present challenges in terms of integrating the test into existing application systems and ensuring that suitable applicants are not inadvertently deterred from applying. If the SJT is to be administered online then appropriate checks should also be put in place to deter cheating. This might include asking applicants to complete the SJT in a proctored setting or implementing a follow-up verification test later in the selection process.

1.21. Feedback from applicants should be gathered to provide insight into their perceptions of the SJT methodology and the suitability of the scenarios. Furthermore, where possible, information regarding the applicant's demographic details should be gathered to inform the analysis of sub-group differences. When gathered, it should be made clear to applicants that this is being used for monitoring purposes only and will not inform the selection decision. Such questions should also be voluntary.

1.22. This review highlights some of the challenges associated with individual Trusts attempting to undertake the development of SJTs in isolation. Such challenges are largely attributed to a lack of resources available locally to invest in development,

refinement, piloting and evaluation in accordance with best practice. An approach whereby a number of Trusts work as a consortium to pool resources and invest in on-going development and evaluation of SJTs may be more likely to result in the development of robust and fit-for-purpose tools.

Appendix v: Case example 3, selection centres

1.1. Three Trusts reported the use of some form of selection centre (SC) to assess values. One was designed for the recruitment of high volume support function roles, one for high stakes senior roles and the other for a broad range of roles across the organisation. The design of SCs generally included the use of a range of tools, which included role play exercises, presentations and group exercises. It should be noted that minimal information was drawn from two examples; therefore the content of this case example is drawn predominately from one Trust using the SC to recruit for high stakes senior roles.

1.2. The case for movement from traditional CV sifts and unstructured interviewing arose from high turnover, low retention, high spends on temporary staff, and poor performance outcomes in a number of roles within the organisation. The Trust was increasingly mindful that the assessment of formal qualifications and past experiences were insufficient to predict performance in a healthcare setting. Initial measures taken to incorporate values throughout the recruitment process involved an in-house development of standalone tools to assess values. These tools were used alongside recruitment events such as open days and induction sessions for the selection of support function roles. Having found favourable results, the organisations sought to adapt the strategy in the recruitment of senior roles using a traditional SC approach.

Developing the selection centre

1.3. Development began with stakeholder consultation to understand the nuances of the role from the job holder's perspective using some form of a job analysis[35] methodology. The programme team recognised that given the seniority of the potential employees and the expected level of experience, there was a lesser need to inform, train and monitor the recruits as part of the selection process. Thus rather than replicating a like-for-like design, the team gave more emphasis to the systematic assessment of values and behaviour.

1.4. The identification of selection methods within the SC was based on the results of the job analysis. Assessment exercises were designed in collaboration with subject matter experts (SMEs), who provided realistic context for the assessment to be based upon. A combination of exercises was used to assess behaviour including a role play exercise, a group discussion exercise, and a scenario-based written exercise or presentation.

1.5. As well as a standard scoring scale for each exercise, the weight given to each assessment score was also based on the outcome of the job analysis. This had informed the values and competencies that were assessed by each exercise and its ratio of redundancy at each stage of the selection processes.

---

[35] A job analysis is a systematic analysis of the relevant knowledge, skills, abilities and attributes associated with performance in the target role. This is the first stage of designing a selection process or method in accordance to best practice.

Implementing the selection centre

1.6. Recruiting managers and members of the programme team were involved as assessors in the implementation of the SC. Prior to the 'live event', assessors, administrators, and role play actors attended training events to understand their role in the SC and what was expected of them. Specifically assessor training events were conducted to help assessors draw behavioural evidence from a candidate's performance, how the evidence may be scored against the scoring structure, and how to go about arriving at a final decision. Emphasis was given to the use of the standard scoring, as its consistency was taken as a measure of fairness across candidates. To enable balanced views, scores were calibrated at the end of the SC and final decisions were jointly made by the recruiting manager and members of programme team.

Evaluating the selection centre

1.7. The effectiveness of the SC underwent various evaluation measures. Stakeholder perceptions of effectiveness were considered alongside objective analyses of workforce metrics, such as changes to retention rates, measured absenteeism, turnover ratio, and average cost per recruit. Adding to these positive results, an evaluation of predictive validity[36] also generated a positive correlation between performance at the SC and subsequent in-role appraisal, suggesting that the results of the SC are likely to predict job performance.

1.8. Trends in relation to sub-group differences were monitored on a year-on-year basis. The work was conducted to monitor the changes in the recruits' profile characteristics to infer possible issues about equality and diversity. So far issues around possible sub-group differences have not been identified.

Next steps

1.9. There is an intention to embed values based recruitment in a greater number of roles through the use of SCs. Whilst strategies have been tested and evaluated in volume support roles and senior level recruits respectively, the next step is to adapt the method further for mid-level roles.

Key learning points and issues for consideration

1.10. The development of an SC should ideally be preceded by a job analysis to identify the values and competencies that are required for success in the role and then mapping such to the NHS Constitution. Results of a job analysis should also inform a set of scoring criteria to guide the evaluation of performance within each individual exercise and the overall SC.

1.11. Stakeholder consultation should provide a role-specific context for which individual exercises and overall SC could be based upon.

---

[36] The predictive validity measures the extent to which outcomes of the selection tool are accurate predictors of subsequent performance; which may be the overall performance of the selection process or post-recruitment in-role performance.

1.12. The overall effectiveness of an SC is reliant on the extent to which assessors are able to extract and score behavioural evidence. It is therefore important that assessors are given adequate training prior to the 'live event'.

1.13. The requirement for adequate assessor training should be balanced with a consideration for the time demand imposed on the recruiting manager.

1.14. The benefits of an SC design may need to be balanced with a consideration of the cost involved to design, implement and evaluate the approach. There is typically greater justification for the recruitment of senior roles using this approach.

1.15. In addition to evaluating the overall outcome, individual selection exercises within an SC may also undergo evaluation of accuracy and effectiveness. Such results may inform not only the effectiveness of the individual exercises, the data collected may also undergo data modelling to identify a fit-for-purpose scoring algorithm and decision making process across an SC.